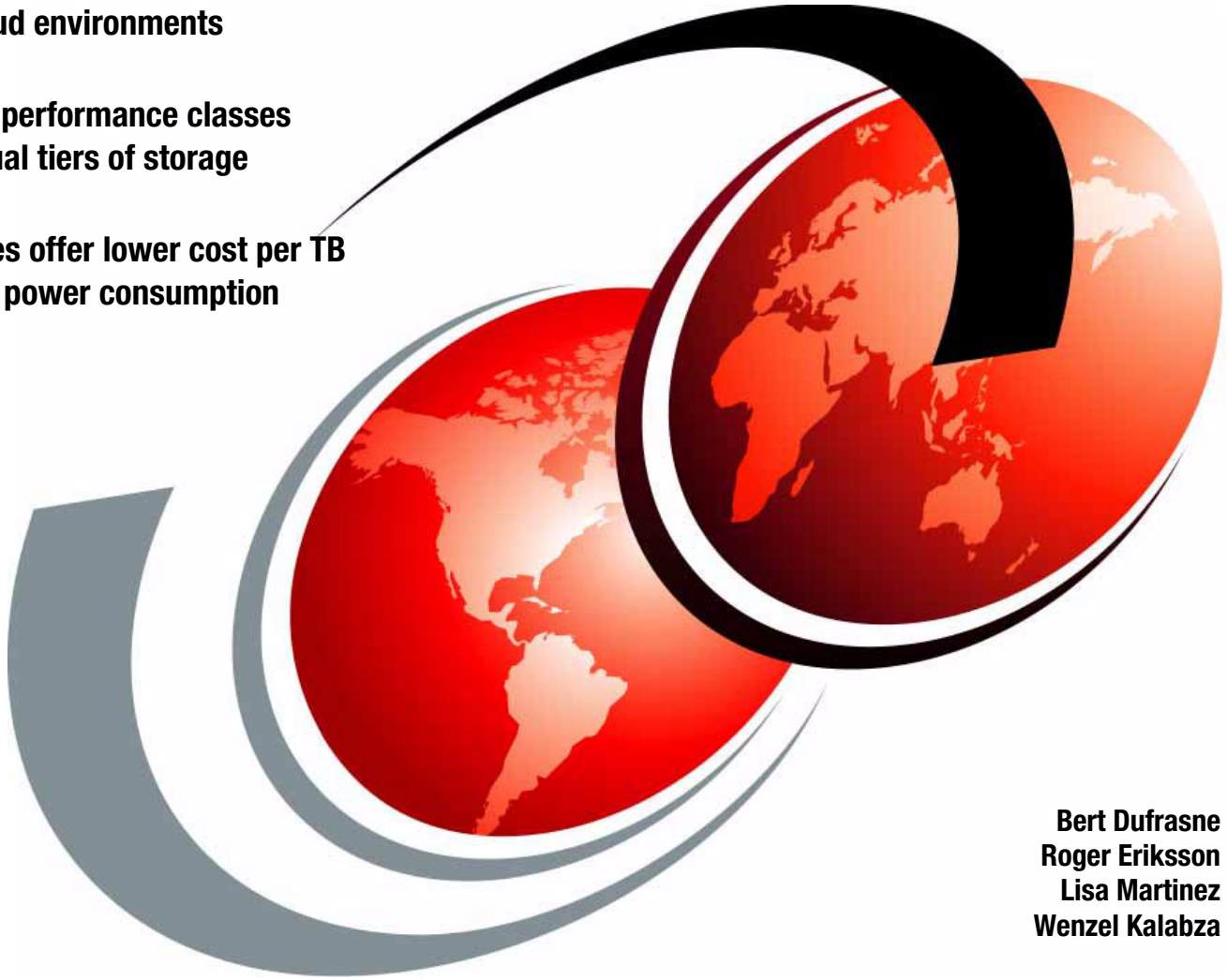
IBM

IBM XIV Storage System Architecture and Implementation

- Multi-tenancy gives more flexibility for XIV in Cloud environments
- Enhanced performance classes allow virtual tiers of storage
- 6 TB drives offer lower cost per TB and lower power consumption



Bert Dufrasne
Roger Eriksson
Lisa Martinez
Wenzel Kalabza

Redbooks

ibm.com/redbooks

International Technical Support Organization

IBM XIV Storage System Architecture and Implementation

May 2014

SG24-7659-08

Note: Before using this information and the product it supports, read the information in “Notices” on page ix.

Ninth Edition (May 2014)

This edition applies to the IBM XIV Storage System with XIV Storage System Software Version 11.5.1 and IBM XIV Storage Management Tools Version 4.4.

© Copyright International Business Machines Corporation 2014. All rights reserved.
Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Notices ix
Trademarks x

Preface xi
Authors xii
Now you can become a published author, too xiii
Comments welcome xiii
Stay connected to IBM Redbooks publications xiii

Summary of changes xv
May 2014, Ninth Edition xv

Chapter 1. IBM XIV Gen3 Storage System overview 1
1.1 Introduction 2
1.2 New features in XIV 3
1.3 Total cost of ownership 4
1.4 XIV Gen3 Storage System models and components 5
1.5 XIV Gen3 design features and functions 6
1.5.1 Massive parallelism 6
1.5.2 Workload balancing 7
1.5.3 Self-healing 7
1.5.4 Fast drive rebuild 7
1.5.5 True virtualization 7
1.5.6 Flash caching (optional) 8
1.5.7 Thin provisioning 8
1.5.8 Processing power 8
1.5.9 SAN connectivity 8
1.5.10 Inter-generational mirroring 8
1.6 XIV Storage System Software 9
1.6.1 IBM XIV Storage Management tools 12
1.7 Host support 15

Chapter 2. IBM XIV Storage System logical architecture and concepts 17
2.1 Architecture overview 18
2.1.1 Features 18
2.1.2 Hardware elements 18
2.2 Parallelism 21
2.2.1 Hardware parallelism and grid architecture 21
2.2.2 Software parallelism 23
2.3 Full storage virtualization 24
2.3.1 XIV Storage System virtualization design 24
2.4 Logical system concepts 26
2.4.1 Logical constructs 26
2.4.2 Logical volume layout on physical disks 28
2.4.3 Multitenancy 30
2.5 System capacity 32
2.5.1 Net usable capacity 33
2.5.2 Global spare capacity 34
2.5.3 Metadata and system reserve 34

© Copyright IBM Corp. 2014. All rights reserved.

2.5.4 Mirrored copies of data 34
2.6 Storage pool concepts 35
2.6.1 Improved management of storage space 35
2.6.2 Consistency groups 35
2.6.3 Storage pool relationships and rules 36
2.7 Capacity allocation and thin provisioning 37
2.8 Flash caching architecture 43
2.8.1 Flash caching overview 43
2.8.2 Flash caching algorithm and concepts 44
2.9 Encryption for data at rest 45
2.10 Reliability, availability, and serviceability (RAS) 46
2.10.1 Resilient architecture 47
2.10.2 Preserving data redundancy: Rebuilding and redistributing 51
2.10.3 Exclusive additional functions for reliability and availability 58

Chapter 3. IBM XIV architecture, components, and planning 61
3.1 IBM XIV Storage System Gen3 62
3.1.1 Fully populated configurations 64
3.1.2 Partially populated configurations 65
3.1.3 Capacity on Demand 66
3.1.4 Advanced System Placement 70
3.1.5 XIV Cloud Storage for Service Providers 72
3.1.6 XIV Storage System Model 114 and Model 214 hardware components 73
3.1.7 Rack 74
3.1.8 Power components 77
3.1.9 Data modules and interface modules 82
3.1.10 Interface modules 87
3.1.11 InfiniBand module interconnect 91
3.1.12 Patch panel 92
3.1.13 Hardware support 93
3.2 Hardware planning overview 94
3.2.1 Basic configuration planning 94
3.2.2 IPv6 addressing and planning 102
3.2.3 Management connectivity 104
3.2.4 IBM XIV Storage System physical installation 105
3.2.5 System power-on and power-off 106

Chapter 4. IBM XIV Storage Management software 111
4.1 XIV Storage Management software introduction 112
4.1.1 XIV Storage Management software platforms 112
4.1.2 XIV Storage Management software interfaces 113
4.2 XIV Storage Management software installation 114
4.2.1 XIV Storage Management (GUI and XCLI) software upgrade 114
4.2.2 XIV Storage Management software installation steps for direct use 115
4.3 XIV Storage Management software use 117
4.3.1 XIV Storage Management GUI used in direct mode 117
4.3.2 XIV Storage System Command-Line Interface 136
4.4 Storage pools 141
4.4.1 Function of storage pools 141
4.4.2 Managing storage pools with the XIV Storage Management GUI 142
4.4.3 Managing storage pools with XIV Command-Line Interface 150
4.5 Volumes 152
4.5.1 Managing volumes with the XIV Storage Management GUI 153

4.5.2 Managing volumes with XIV Command-Line Interface 162
4.6 Host definition and mapping 164
4.6.1 Assigning LUNs to a host using the XIV Storage Management GUI 164
4.6.2 Assigning logical unit numbers to a host by using the XCLI 168
4.7 QoS feature 170
4.7.1 Managing QoS 171
4.8 Multitenancy 174
4.8.1 Domains 175
4.8.2 Managing domains with the XIV Storage Management GUI. 175
4.8.3 Limiting domains by performance class (QoS) 193
4.8.4 Domains and LDAP authentication 195
4.8.5 Managing domains with the XIV command-line interface, XCLI 197
4.9 Flash cache 199
4.9.1 Managing flash cache with the XIV Storage Management GUI 199
4.9.2 Managing flash cache with XIV Command Line Interface. 204

Chapter 5. Security 207
5.1 Physical access security 208
5.2 x509 certificate validation and management 208
5.2.1 Managing x509 certificates with the XIV Storage Management GUI 209
5.3 Configuring IPv6 addresses 215
5.3.1 Enabling IPv6 support 215
5.3.2 Configuring IPv6 addresses 215
5.4 Configuring Internet Protocol Security connectivity. 217
5.5 Native user authentication. 219
5.5.1 XIV Storage System local credential repository 220
5.5.2 Managing user accounts with the XIV Storage Management GUI 225
5.5.3 Security considerations when using Hyper-Scale Manager 232
5.5.4 Managing user accounts using the XCLI 234
5.5.5 Password management and resets. 237
5.5.6 Managing multiple systems. 238
5.6 Enabling compliance with PCI-DSS 241
5.6.1 Auditing via syslog 242
5.6.2 Define Idle Timeout to lock a GUI or XCLI session. 243
5.7 LDAP-based authentication 244
5.7.1 Introduction to Lightweight Directory Access Protocol 244
5.7.2 LDAP directory components 245
5.7.3 LDAP product selection 246
5.7.4 LDAP role mapping. 247
5.8 Defining LDAP on the XIV Storage System 251
5.8.1 Using the XIV Storage Management GUI LDAP wizard to configure LDAP 252
5.8.2 Using the XIV Storage Management GUI directly to configure LDAP. 259
5.8.3 Using XIV Storage System command-line interface 262
5.9 LDAP-managed user authentication 263
5.9.1 LDAP repository credential objects 264
5.9.2 Managing LDAP user accounts 268
5.9.3 Managing user groups by using the XCLI in LDAP authentication mode 274
5.9.4 Active Directory group membership and XIV role mapping 274
5.9.5 OpenLDAP Directory and XIV Storage System role mapping 278
5.9.6 Managing multiple systems in LDAP authentication mode 282
5.10 Securing LDAP communication with Secure Sockets Layer 283
5.10.1 Configuring XIV to use LDAP over Secure Sockets Layer 283
5.10.2 Maintaining the SSL certificate 285

Chapter 6. Performance 287
6.1 XIV Storage System Software and hardware architecture 288
6.1.1 Workload distribution and load balancing 288
6.1.2 Grid architecture 290
6.1.3 Caching mechanisms 291
6.1.4 Data redistribution effects on host systems 292
6.1.5 Snapshots 293
6.2 Practices for optimum performance 294
6.2.1 Sizing 294
6.2.2 Number of logical unit numbers 295
6.2.3 Multipathing considerations 297
6.2.4 Host considerations 299
6.2.5 Quality of service 301
6.3 Performance monitoring 304
6.3.1 Using the XIV Storage Management GUI 304
6.3.2 Using the XIV Top utility 312
6.3.3 Using the XIV Storage System command-line interface 316
6.3.4 Tivoli Storage Productivity Center 318
6.4 Performance evaluation 320
6.4.1 Problem-solving steps 320

Chapter 7. Monitoring 327
7.1 Monitoring with XIV Storage Management GUI and XCLI 328
7.1.1 Monitoring by using the Systems menu 328
7.1.2 Monitoring alerts for all defined IBM XIV systems with the GUI 330
7.1.3 Monitoring an individual XIV Storage System using the GUI 331
7.1.4 Monitoring with XIV Storage System Command-Line Interface 339
7.1.5 XIV audit event logging 345
7.1.6 Viewing events in the XIV Storage Management GUI 346
7.1.7 Event attributes 346
7.1.8 Viewing events using the XCLI 347
7.1.9 Defining notification rules 350
7.2 Monitoring using the IBM XIV Mobile Dashboard 351
7.2.1 Installing the IBM XIV Mobile Dashboard 351
7.2.2 Mobile Dashboard for Android devices 356
7.2.3 IBM Storage Mobile dashboard 357
7.3 XIV Storage System event notification 359
7.3.1 Configuring the XIV Storage System to send notifications 359
7.3.2 Setting up event notification 359
7.4 Mobile push notification 368
7.5 Call Home and remote support 371
7.5.1 Call Home feature 372
7.5.2 Remote support 372
7.5.3 XIV Remote Support Proxy 380
7.5.4 Installation 381
7.6 Simple Network Management Protocol-based monitoring 384
7.6.1 Monitoring tools 384
7.6.2 Using SNMP commands to confirm the XIV Storage System status 386
7.6.3 Using SNMP get or walk commands with open source software 387
7.7 Using Tivoli Storage Productivity Center 390
7.7.1 Setting up and discovering XIV systems in Tivoli Storage Productivity Center . . . 390
7.7.2 XIV Storage System Tivoli Storage Productivity Center reports 396
7.7.3 Tivoli Storage Productivity Center web-based GUI 404

7.8 Custom monitoring 407
7.8.1 Custom monitoring of a volume 407
7.8.2 Custom monitoring of performance attributes 410
7.8.3 Summary 412
7.9 Microsoft System Center Operations Manager 413
7.9.1 Prerequisites 413
7.9.2 Installing SCOM and the IBM Storage Management Pack 414
7.9.3 Importing the management packs and Adding IBM XIV systems 415
7.9.4 Configuring the SCOM Management pack 418
7.9.5 Monitoring your XIV Storage System with SCOM 419
7.9.6 Upgrading the IBM Storage Management Pack 423

Appendix A. Thin provisioning conceptual examples 425
System-level thin provisioning conceptual example 426
Regular storage pool conceptual example 427
Thinly provisioned storage pool conceptual example 428

Appendix B. Additional LDAP information 431
Creating user accounts in Microsoft Active Directory 432
Creating user accounts in Oracle Java Directory 438
Securing LDAP communication with SSL 447
Windows Server SSL configuration 447
Oracle Java Directory SSL configuration 454
Certificate authority setup 460

Related publications 465
IBM Redbooks publications 465
Other publications 465
Online resources 465
How to get IBM Redbooks publications 466
Help from IBM 466

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

© Copyright IBM Corp. 2014. All rights reserved.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®
DS6000™
DS8000®
FlashSystem™
IBM®
IBM FlashSystem™
IBM Flex System®
NetView®
POWER®
Redbooks®
Redpaper™
Redbooks (logo) ®
Storwize®
System p®
System Storage®
Tivoli®
XIV®

The following terms are trademarks of other companies:

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Other company, product, or service names may be trademarks or service marks of others.

Preface

This IBM® Redbooks® publication describes the concepts, architecture, and implementation of the IBM XIV® Storage System.

The XIV Storage System is a scalable enterprise storage system that is based on a grid array of hardware components. It can attach to both Fibre Channel Protocol (FCP) and IP network Small Computer System Interface (iSCSI) capable hosts. This system is a good fit for clients who want to be able to grow capacity without managing multiple tiers of storage. The XIV Storage System is suited for mixed or random access workloads, including online transaction processing, video streamings, images, email, and emerging workload areas, such as Web 2.0 and cloud storage.

The focus of this edition is on the XIV Gen3 running Version 11.5.x of the XIV system software, which brings enhanced value for the XIV Storage System in cloud environments. It offers multitenancy support, VMware vCloud Suite integration, more discrete performance classes, and RESTful API enhancements that expand cloud automation integration. Version 11.5 introduces support for three-site mirroring to provide high availability and disaster recovery. It also enables capacity planning through the Hyper-Scale Manager, mobile push notifications for real-time alerts, and enhanced security. Version 11.5.1 supports 6TB drives and VMware vSphere Virtual Volumes (VVOL).

In the first few chapters of this book, we describe many of the unique and powerful concepts that form the basis of the XIV Storage System logical and physical architecture. We explain how the system eliminates direct dependencies between the hardware elements and the software that governs the system. In subsequent chapters, we explain the planning and preparation tasks that are required to deploy the system in your environment by using the intuitive yet powerful XIV Storage Manager GUI or the XIV command-line interface. We also describe the performance characteristics of the XIV Storage System and present options for alerting and monitoring, including enhanced secure remote support.

This book is for IT professionals who want an understanding of the XIV Storage System. It is also for readers who need detailed advice on how to configure and use the system.

Copy services and data migration features are covered in the Redbooks publication, *IBM XIV Storage System Business Continuity Functions,* SG24-7759.

Host operating systems and other integration aspects are addressed in a separate publication, *XIV Storage System: Host Attachment and Interoperability*, SG24-7904. Refer also to the IBM Redpaper™ publication, *Using XIV in VMware environments*, REDP-4965. For details about thin provisioning and space reclamation, see the Redpaper, *XIV Thin Provisioning and Space Reclamation,* REDP-5001.

For information about IBM Hyper-Scale, see the Redpaper, *IBM Hyper-Scale for the XIV Storage System*, REDP-5053. For information about encryption, see the Redpaper, *XIV Security with Data-at-Rest Encryption*, REDP-5047.

© Copyright IBM Corp. 2014. All rights reserved.

Authors

This book was produced by a team of specialists from around the world working for the International Technical Support Organization, at the IBM European Storage Competence Center in Mainz, Germany.

Bertrand Dufrasne is an IBM Certified Consulting I/T Specialist and Project Leader for IBM System Storage® disk products at the International Technical Support Organization (ITSO), San Jose Center. He has worked at IBM in various I/T areas. He has authored many IBM Redbooks publications, and has also developed and taught technical workshops. Before joining the ITSO, he worked for IBM Global Services as an Application Architect. He holds a Master's degree in Electrical Engineering.

Roger Eriksson is an STG Lab Services consultant, based in Stockholm, Sweden, who works for the European Storage Competence Center in Mainz, Germany. He is a Senior Accredited IBM Product Service Professional. Roger has over 20 years of experience working on IBM servers and storage, including Enterprise and Midrange disk, NAS, SAN, IBM System x, IBM System p®, and IBM BladeCenter. He has done consulting, proof of concepts, and education, mainly with the XIV product line, since December 2008. He has worked with both clients and various IBM teams worldwide. He holds a Technical College Graduation in Mechanical Engineering.

Lisa Martinez has been working in the North America Storage Specialty Team (formerly ATS) as a storage consultant since January 2012. Her focus has been with pre-sales support for DS8000® and XiV as well as lead instructor for XiV customer based workshops. Prior experience includes roles as a storage architect in the Specialty Services Area in GTS, a temporary assignment as a Global Support Manager for Cardinal Health and test architect in disk storage focusing on system level test for XiV for three years, and copy services for DS8K. Lisa holds degrees in Computer Science from New Mexico Highlands University and Electrical Engineering from the University of New Mexico. She has been employed with IBM for 17 years.

Wenzel Kalabza s a certified XIV Product Field Engineer (PFE) based in the storage competence center in Mainz, Germany. Wenzel joined IBM in 1998 as customer quality engineer for IBM disk drive failure and performance analysis. He joined the Back Office for the high end storage system (ESS) in June 2002. In 2005, Wenzel started a PFE role for the IBM disk storage DS6000™. In June 2008, he became a PFE for the XIV storage product. Wenzel holds a degree in Electrical Engineering and Power Economy, and several storage related certifications.

Thanks to the following people for their contributions to this project:

Eyal Abraham, Diane Benjuya, Ramy Buechler, Rami Elron, Theodore Gregg, Peter Kisich, Rony Shapiro, Yossi Siles, Oded Kellner, George Thomas, Carlo Saba, Ohad Atia, Marcus Boelke, Daniel Lereya, Mary J. Connell.
IBM

Thanks also to the authors of the previous editions:

Jana Jamsek, Suad Musovich, Markus Oscheka, In Kyu Park, Francesco Perillo, Carlo Saba, Hank Sautter, Jim Sedgwick, Eugene Tsypin, Kip Wagner, Alexander Warmuth, Peter Wendler, Axel Westphal, Ralf Wohlfarth, Dietmar Dausner, Itzhack Goldberg, Stephen Solewin, Christian Schoessler, Patrick Schill, Christian Burns, Thomas Peralto.

Special thanks to ESCC team in IBM Mainz, Germany, for hosting the project and making equipment available in their lab.

Now you can become a published author, too

Here's an opportunity to spotlight your skills, grow your career, and become a published author, all at the same time. Join an ITSO residency project and help write a book on your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us.

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** IBM Redbooks publications form:

  ibm.com/redbooks

- Send your comments by email:

  redbooks@us.ibm.com

- Mail your comments:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks publications

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks publications weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

Summary of changes

This section describes the technical changes made in this edition of the book and in previous editions. This edition might also include minor corrections and editorial changes that are not identified.

Summary of Changes
for SG24-7659-08
for IBM XIV Storage System Architecture and Implementation
as created or updated on May 13, 2015

May 2014, Ninth Edition

This revision reflects the addition, deletion, or modification of new and changed information described below.

New information

- Inclusion of XIV Storage System Software Version 11.5 features, including multitenancy, enhanced performance classes, mobile notifications, and security updates
- Examples and illustrations that reflect XIV Storage Management GUI v4.4
- Updated, “Similar results can be accomplished on UNIX-based platforms by using shell scripting and crontab for scheduling custom monitoring processes.” on page 412

Changed information

- Various updates to reflect the XIV Storage Software Version 11.5.1 features
- IBM XIV Gen3 latest hardware

© Copyright IBM Corp. 2014. All rights reserved.

1

IBM XIV Gen3 Storage System overview

The IBM XIV Gen3 Storage System is a high-end disk storage system. It provides secure, self-service, scalable, enterprise-class storage with a self-healing, data-protected cloud-centric foundation that is ideal for public, private, or hybrid clouds. The XIV cloud economics provide world-class business continuity for applications that require zero tolerance for downtime. It's predictable, consistent, high-performance design enables organizations to take control of their storage and gain business insights from their data. The storage system contains proprietary and innovative algorithms that offset hardware malfunctions, minimize maintenance, provide flexibility, and drive massive parallelism. XIV provides a high service level for dynamic and mixed workloads.

This chapter provides a high-level overview of the XIV Gen3. It covers the following topics:

- 1.1, "Introduction" on page 2
- 1.2, "New features in XIV" on page 3
- 1.3, "Total cost of ownership" on page 4
- 1.4, "XIV Gen3 Storage System models and components" on page 5
- 1.5, "XIV Gen3 design features and functions" on page 6
- 1.6, "XIV Storage System Software" on page 9
- 1.7, "Host support" on page 15

© Copyright IBM Corp. 2014. All rights reserved.

1.1 Introduction

The XIV Gen3 Storage System architecture delivers performance, scalability, and ease of management while harnessing the high capacity and cost benefits of serial-attached SCSI (SAS) drives. The system uses off-the-shelf components, therefore allowing for faster adoption of newer generation hardware.

The XIV Storage System incorporates fully automated recovery processes. A data redistribution takes place automatically after new hardware is added, removed, or has failed. Only the data necessary to keep the system fully redundant is redistributed, minimizing recovery time and data movement. Because of this powerful distribution algorithm, the performance is always consistent.

The virtualized grid architecture and algorithms that are used divide the host data into 1 MB partitions and distribute them pseudo-randomly across all disks, leading to a consistent load on all components and eliminating hotspots. During the distribution process, the data is always mirrored by ensuring that each 1 MB partition is kept in at least two separate locations within the system.

Clients receive exceptionally low total cost of ownership because the system software licensing includes quality of service (QoS), snapshots, thin provisioning, data migration, asynchronous and synchronous mirroring, and an intuitive GUI, combined with dramatic efficiencies in capacity, power, and space. Because of the unique physical architectural design of the system, including off-the-shelf modules, network switches, and power components, new technologies can be adopted easily. XIV offers encryption for data at rest while avoiding performance impact

Flash (solid-state drive or SSD) caching, available as an option, adds up to 12 TB[1] of management-free caching power to the whole system, providing up to four times better performance for application workloads, without the need for setup, administration, or migration policies.

The IBM XIV Gen3 Storage System is fully virtualized storage designed to eliminate the need for performance tuning, planning for capacity and performance growth, and numerous other storage management activities. The highly intuitive XIV GUI and built-in management tools make administrative tasks easy and efficient, with little training or expertise required, from provisioning volumes to monitoring multiple systems. A powerful command-line interface (CLI) supports complex scripting. The unified console enables one-stop centralized administration of multiple XIV systems. Its exceptional flexibility extends to mobile devices, giving users the flexibility of performance and capacity monitoring by the IBM XIV Mobile Dashboard, supporting the Apple iPhone and Apple iPad, and Android devices.

These design points provide the following benefits:

- Reduced total cost of ownership (TCO)
- Flexibility to meet changing business needs
- Consolidated and virtualized data storage
- High availability and performance characteristics.

[1] Twelve TB using 15 x 800 GB flash drives in XIV Gen3 systems equipped with 4 TB or 6 TB drives

1.2 New features in XIV

The XIV System Software version 11.5.0 and 11.5.1 and other recent offerings provide a range of new capabilities that enhance its value for cloud services:

- IBM XIV Cloud Storage for Service Providers, introduced in May 2014, offers a pay-per-need scalability matrix. This enables cloud providers to cater to tenant requirements with ease and efficiency through flexible software licensing. The portfolio capitalizes on the v11.5.0 enhancements and multitenancy isolation, which enables more than 120 tenants to securely access and manage data while sharing the same XIV Storage System infrastructure.
- XIV multitenancy securely isolates domains of XIV storage resources among numerous tenants, with the ability to set different quality of service (QoS) levels for each domain. It enables the division of storage system administration tasks into logical domains, using role-based permissions. It also enables rapid deployments while minimizing the need for extensive planning, tuning, or field upgrades.
- Support for RESTful API version 2.0 adds functions for storage administration tasks, including Copy Services.
- XIV and VMware vCloud Suite Integration functions are available to all XIV releases through the IBM Storage Integration Server v1.5. It is free of charge and enables VMware APIs for Storage Awareness (VASA), VMware vSphere Web Client, vCenter Orchestrator (vCO), vCenter Operations Manager (vCOPs), and VMware vCloud Automation Center (vCAC). As such, the IBM Storage Integration Server delivers a range of IBM storage integration services in cloud based architectures providing provisioning, automation, and monitoring.
- Three-site mirroring, also referred to as 3-way mirroring, allows a 3-way replication scheme where data is mirrored among three mirroring sites (peers) using one synchronous and one asynchronous connection in a concurrent topology. Thanks to XIV's unique architecture, the multiple mirroring has near-zero impact on system performance.

  The function is only available on Gen3 systems (with XIV Storage Software v11.5 or later). There is no additional cost, or license requirement.
- The IBM Hyper-Scale Manager now includes a reporting feature that graphically depicts trending of capacity and use across multi-system deployments. It's a built-in feature that does not require any add-ons, tools, or external databases.
- GUI support for the cross-system consistency groups in IBM Hyper-Scale Manager that automates the coordinated creation of snapshots for interdependent consistency groups on multiple XIV systems.
- Security enhancements including auditing of user actions and user interface locking after a predefined period of idle time, allow XIV to meet requirements imposed by the Payment Card Industry Data Security Standards (PCI-DSS).
- Mobile Push notifications, enables XIV clients to receive high severity events on mobile devices (iPhone, Android platforms), providing configurable real-time alerts, anytime, anywhere.
- The quality of service (QoS) now offers up to 500 performance classes and is controllable on a host, or domain and pool-based level.
- Capacity alert thresholds can now be set on a per pool (and per domain) basis.
- The XIV GUI is getting more global and now supports three languages: English, Japanese, and Chinese.

Version 11.5.1 introduces:

- VMware Virtual Volumes (VVol) , enabling upcoming VMware vSphere environments to easily automate XIV provisioning, offload snapshots and cloning, and instantly reclaim space, all at the virtual-machine level and while benefiting from XIV. VMware VVOLs automation is based on VMware vSphere APIs for Storage Awareness (VASA). The IBM Storage Provider for VMware VASA is a feature of the IBM Storage Integration Server, and will support the orchestration of all Vitual Volumes operations with XIV. Support for VVOL will be provided through IBM Storage Integration Server v2.0.
- Storage Management Initiative Specification (SMI-S) 1.6 latest protocol certification helps build dynamic, scalable, secure Microsoft-based cloud storage infrastructures with out-of-the box integration with Microsoft System Center Virtual Machine Manager 2012.
- 6 TB Self Encrypting Drives (SED).The XIV with 6 TB drives lets you store 489TB of data in a single rack using the same floor space, but with a lower cost per TB and lower power consumption per TB than the previous systems. Compared to XIV equipped with 4 TB drives, the 6TB drives offer 33 percent lower power consumption per TB at up to 12 percent lower price.

1.3 Total cost of ownership

The total cost of ownership (TCO) calculation for traditional enterprise storage systems usually includes the following costs:

- Cost of initial purchase
- Cost of subsequent upgrades
- Cost of hardware and software maintenance
- Cost of storage management
- Cost of electricity and cooling

The XIV Storage System includes software licenses for all features and functions at no additional charge. There is no need to purchase additional software licensing when you decide to add more capacity to your system or use an advanced feature, such as mirroring or data migration. To augment the capacity, additional hardware must be purchased. Clients may also take advantage of the Capacity on Demand option, acquire XIV systems using Advanced System Placement, or capitalize on IBM XIV Cloud Storage for Service Providers solution offering.

The IBM XIV Storage System Management software suite, along with the virtualized grid architecture, greatly simplifies the layout and management of data within the storage system, which reduces the cost of managing it. Snapshots and test environments are created in seconds. Data migration is dramatically simple and fast; remote mirroring is easy and is supported even between XIV Gen3 and the XIV second-generation Storage System. By reducing complexity, the system minimizes the IT resources required to manage storage, freeing individuals for other tasks.

The XIV Storage System Gen3 uses a grid array of low cost, high capacity ([2]2 TB, 3 TB, 4 TB, or 6 TB) SAS drives for storing data, which provides performance similar to Fibre Channel (FC) drives in traditional storage systems. The grid architecture used in the XIV Storage System eliminates the need for “idle spare drives” in the event of a drive failure. As a result, all drives in the system are fully used, reducing the number of idle components.

[2] The 1 TB drives are actually capacity limited 2 TB drives.

The system's use of very high-density drives maximizes physical space per floor tile, offering maximum capacities of 489 usable TB (in 6 TB disk drive configurations). High-density storage can significantly reduce power, space, and cooling costs, and offers much lower cost per terabyte relative to comparable systems.

These features and more reduce the overall TCO of the XIV Storage System.

1.4 XIV Gen3 Storage System models and components

The XIV Gen3 Storage System family consists of two machine types and two models:

- Machine type-model 2812-114, 2812-214
- Machine type-model 2810-114, 2810-214

Note: Effective October 25, 2013, the XIV Gen3 2812-114 and 2810-114 have been withdranw from marketing and can no longer be ordered from IBM.

The 2812 supports a 3-year warranty to complement the 1-year warranty offered by the existing and functionally equivalent 2810.

All of the machine types are available in the following modules' configurations:

- Six modules (including three Interface Modules)
- Nine to fifteen modules (including six Interface Modules)

The 114 model includes the following components, which are visible in Figure 1-1

- Three to six Interface Modules, each with 12 SAS disk drives (2 TB or 3 TB but no intermixing).
- Three to nine Data Modules, each with 12 SAS disk drives (2 TB or 3 TB but no intermixing).
- Flash caching support. Each Data or Interface Module can be equipped with one 400 GB flash drive (SSD) as fast read cache (6 TB for a full system with 15 modules).
- An uninterruptible power supply (UPS) module complex comprising three redundant UPS units.
- Two InfiniBand module interconnects with redundant power supplies (RPSs).
- A Maintenance Module.
- An Automatic Transfer Switch (ATS) for external power supply redundancy.
- A modem, which is connected to the maintenance module for external system service.

The model 214 includes the same components and brings the following enhancements:

- Data modules and interface modules can be equipped with 2 TB, 3 TB, 4 TB, or 6 TB, but no intermixing
- Up to twelve 10 GbE ports for connecting to iSCSI-attached hosts (or twenty-two 1 GbE ports, as in the 114 model)
- Up to 15 CPUs providing 90 physical cores (180 logical cores using Intel Hyper-Threading technology)
- More energy-efficient hardware that can reduce power consumption by up to 16% compared to previous models

- In configurations using the 4 TB or 6 TB drives, an 800 GB flash drive is available, increasing the read cache to 12 TB for a full system with 15 modules.

For both models (114 and 214), all of the modules in the system are linked through the two internal redundant InfiniBand module interconnects, which enable maximum bandwidth use and are resilient to at least a single component failure.

The system (models 114 and 214) and all of its components come pre-assembled and wired in a lockable rack.

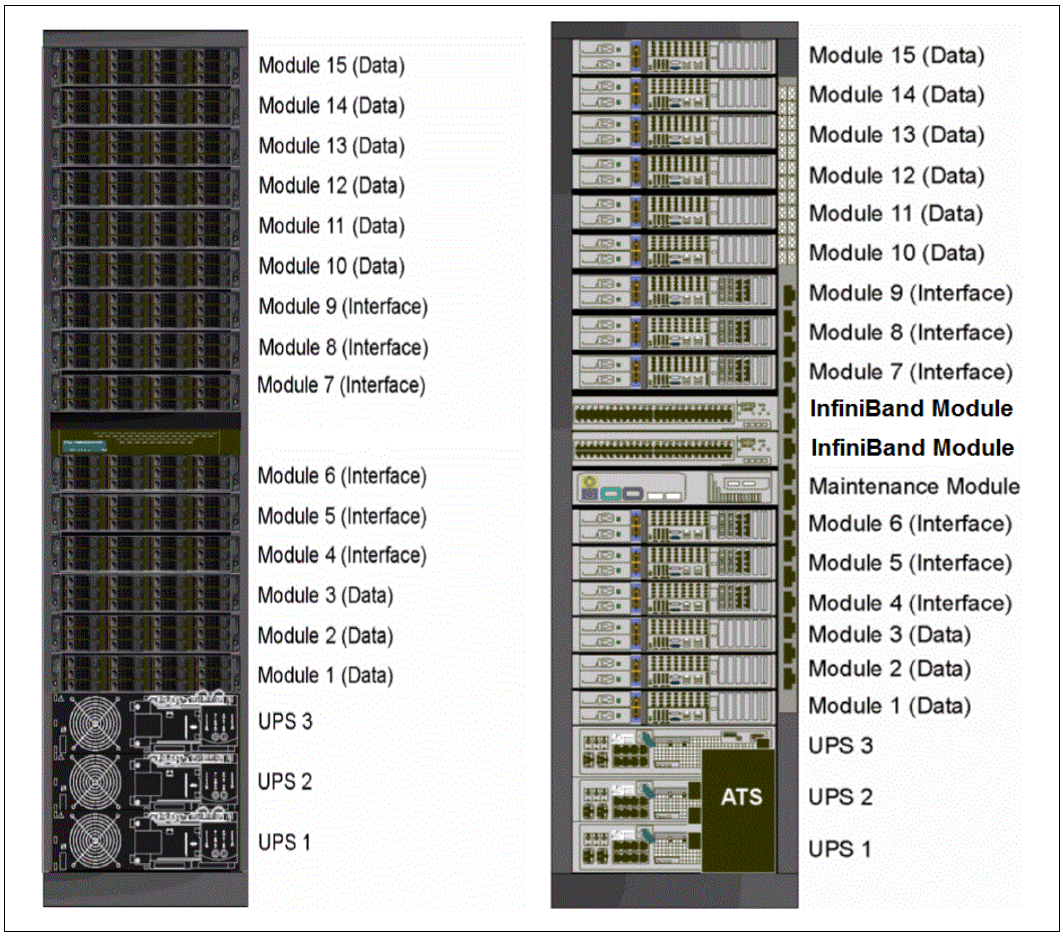


Figure 1-1 XIV Storage System components: Front and rear views

1.5 XIV Gen3 design features and functions

This section describes the key design features and functions of the XIV Storage System architecture.

We describe these key design points and underlying architectural concepts in detail in Chapter 2, “IBM XIV Storage System logical architecture and concepts” on page 17.

1.5.1 Massive parallelism

The system architecture ensures full use of all system components. Any input/output (I/O) activity involving a specific logical volume in the system is always inherently handled by all

spindles. The system harnesses all storage capacity and all internal bandwidth. It also takes advantage of all available processing power for host-initiated I/O activity and system-initiated activity, such as rebuild processes and snapshot generation. All disks, processors, switches, and other components of the system contribute to the performance of the system at all times.

1.5.2 Workload balancing

The workload is evenly distributed over all hardware components at all times. All disks and modules are used equally, regardless of access patterns. Although applications might access certain volumes more frequently than other volumes, or access certain parts of a volume more frequently than other parts, the overall load on the disks and modules is balanced perfectly.

Pseudo-random distribution ensures consistent load-balancing even after adding, deleting, or resizing volumes, and adding or removing hardware. This balancing of all data on all system components minimizes the possibility of a hotspot.

1.5.3 Self-healing

Protection against concurrent double disk failure is provided by an efficient rebuild process that brings the system back to full redundancy in minutes. In addition, the XIV Storage System extends the self-healing concept, resuming redundancy even after failures in components other than disks, such as a failure of a whole module.

1.5.4 Fast drive rebuild

The enhanced XIV architecture improves drive rebuild times and remains consistent at around 13 minutes per TB of data on a 15-module system. XIV does not rebuild unused (zeroed) blocks.

Important: Identified rebuild times are estimated and vary in specific environments. You might experience reduced or increased rebuild times,depending on architecture, block data written, and resource allocations of modules. Refer to 2.10.2, “Preserving data redundancy: Rebuilding and redistributing” on page 51.

1.5.5 True virtualization

Unlike other system architectures, storage virtualization is inherent in the basic principles of the XIV Storage System design. Physical drives and their locations are hidden from the user, which dramatically simplifies storage configuration and lets the system lay out the user’s volume in the optimal way. The automatic layout maximizes the system’s performance by using system resources for each volume, regardless of the user’s access patterns. With the XIV Storage System design, you no longer must decide which type of data protection to use for storing data or how many drives you need to dedicate for a specific application. Allocate the needed storage space and the system does the rest. This feature allows the user to respond to growing storage needs in minutes instead of hours or days, as with traditional storage systems.

1.5.6 Flash caching (optional)

For ultra-high performance needs, XIV Gen3 optionally offers up to 12 TB of management-free flash caching, which is available to all system data and can be installed nondisruptively. Operating with advanced flash algorithms, the latest XIV Gen3 components help meet requirements for extremely high performance workloads. The latest Gen3 model can provide over its predecessors, up to 4.5 times better performance for random database-type workloads and up to 13.7 GBps sequential read throughput.

1.5.7 Thin provisioning

The system supports thin provisioning at the storage pool level, which is the capability to allocate additional storage space to applications on a just-in-time and as-needed basis, allowing the most efficient use of available resource. As a result, significant cost savings are achieved, compared to traditional provisioning techniques. This benefit is achieved by defining a logical capacity that is larger than the actual physical capacity used.

1.5.8 Processing power

The XIV Storage System open architecture uses the latest processor technologies and is more scalable than solutions that are based on a closed architecture.

1.5.9 SAN connectivity

The XIV Storage System provides up to 24 x 8 Gb FC ports and up to 12 x 10 Gb Ethernet (or 22 x 1 Gb Ethernet) ports for iSCSI connection.

1.5.10 Inter-generational mirroring

Mirroring between XIV Gen3 (model 114 or model 214) and XIV second generation (model A14) systems is supported.

The following prerequisites must be met, however:

- The XIV Storage System Software Version 11.2.0 or later is required for the Gen3 systems Models 114, and Version 11.2.0 or later is required for the Gen3 Model 214.
- The XIV Storage System Software Version 10.2.4e or later is required on the second-generation systems.

The mirroring setup and functions are otherwise unchanged and fully supported between the two generations. For details about the synchronous and asynchronous mirroring functions, see the IBM Redbooks publication titled *IBM XIV Storage System Business Continuity Functions*, SG24-7759.

IBM XIV three-way mirroring that was introduced in the 11.5.0 system software further enhances the system high availability and disaster recovery options.

3-way mirroring: XIV Gen3 Model running system software version 11.5.0 or later is required in supporting three-way replication feature. (Gen2 models are not supported)

1.6 XIV Storage System Software

XIV Storage System Software Version 11.5.x, in conjunction with XIV Gen3 systems, provides the functions of the system, which include the following features:

- Bundled advanced features:

  All the features of the XIV Storage System, including advanced features, such as migration and mirroring, are included at no additional charge and apply to the entire storage capacity.

- Non-Disruptive Code Load:

  System software code can be upgraded without requiring downtime. This feature enables “non-stop” production environments to remain running while new code is upgraded.

  The code upgrade is run on all modules in parallel and the process is fast enough to minimize the impact on host applications.

  No data migration or rebuild processing is allowed during the upgrade. Mirroring, if any, is suspended during the upgrade and automatically reactivated upon completion.

  Storage management operations are also not allowed during the upgrade, although the status of the system and upgrade progress can be queried. It is also possible to cancel the upgrade process up to the point of no return.

- Support for flash (SSD) caching:

  The XIV software manages the flash caching. There is nothing that the storage administrator must configure. The storage administrator can enable or disable the extended flash cache at the system level or on a per host volume level. The XIV software will dynamically and adaptively use the flash as an extended read cache to boost application performance.

  Also, by storing and computing all flash cache data integrity checksum activities in dynamic random access memory (DRAM) during normal operation, rather than on the flash, flash caching can increase performance by up to four times for active application workloads.

  SSD flash caching: XIV offers extensive I/O analytic trace collections from the block storage system level that produces visualization and very good performance benefit estimates for using SSD flash caching. The trace can be captured only by your local IBM Storage Solution Architect. It can help with determining whether SSD flash caching is a cost-effective feature for your XIV infrastructure.

- XIV multitenancy:

  Enables the division of storage system management scheme into logical domains, with extreme flexibility in assigning policy, administrators, and QoS per domain.

- Support for multiple snapshots:

  The snapshot capabilities within the XIV Storage System Software use a metadata, redirect-on-write design that allows snapshots to occur in a subsecond time frame with little performance impact. The system supports multiple differential snapshots of a volume. Any of the snapshots can be made writable, and then snapshots can be taken of the newly writable snapshots (snapshots of snapshots). Volumes can even be restored from these writable snapshots.

- Cross-system consistency groups

These groups enable coordinated creation of snapshots for interdependentgroups on multiple XIV systems. This function is now automated by the GUI in the Hyper-Scale Manager.

- Synchronous and asynchronous remote mirroring to another XIV Storage System:

  Synchronous or asynchronous remote mirroring can be performed over FibreChannel (FC) or IP (iSCSI) connections. Both protocols are also supported for three-way mirroring connectivity. Synchronous remote mirroring is used when a zero recovery point objective (RPO) is required. For practical reasons (latency), ensure that the distance is less than 100 km (62 miles). For longer distances, asynchronous replication is more appropriate.

  Version 11.5.x of the XIV software offers quick-start synchronous mirroring with offline initialization and flexible switching between synchronous and asynchronous mirroring.

  XIV 3-way mirror: XIV software v11.5.x supports a three-site mirroring star topology with one synchronous and one asynchronous mirror relationship. For details, see the IBM Redpaper publication titled *IBM XIV Storage System Multi-site Mirroring*, REDP-5129 and the IBM Redbooks publication titled *IBM XIV Storage System Business Continuity Functions*, SG24-7759.

- Support for thin provisioning:

  Thin provisioning allows administrators to overprovision allocated storage within a storage pool. This task is done by defining volume sizes that are larger than the physical capacity of the storage pool. Unlike other approaches, the physical capacity of the storage pool needs to be only larger than the actual size of the volumes used within that pool. Physical capacity of the storage pool needs to be increased only when actual written data of the volumes comes close to the physical pool size.

- Support for Microsoft Windows Server 2012:

  The XIV Storage System enables you to take advantage of recent Microsoft Windows enhancements, such as space reclamation, that can help you increase capacity use and lower TCO.

- Support for in-band data migration of heterogeneous storage:

  The XIV Storage System is also capable of acting as a host, gaining access to volumes on an existing storage system. The system is configured as a proxy to respond to requests between the current hosts and the storage while migrating all existing data in the background. In addition, the XIV Storage System supports thick-to-thin data migration, which allows the system to reclaim any allocated space that is not occupied by actual data.

- Authentication using Lightweight Directory Access Protocol (LDAP):

  LDAP can be used to provide user logon authentication, allowing the XIV Storage System to integrate with Microsoft Active Directory, Open LDAP, or Oracle Java Systems Directory Server. Multiple directory servers can be configured to provide redundancy in case one server becomes unavailable. LDAP support continues within the recent release of multitenancy implementations offered in v11.5.x or later system code.

  LDAP with multitenancy: One limitation is a similar/duplicate/identical user assigned in two separate domains will not be able to have different user credentials within the same XIV system using LDAP in a multitenancy implementation.

- Encryption for data-at-rest:

IBM XIV Storage System Gen3 helps secure data with industry-standard encryption for data at rest while avoiding performance impact.

For details about encryption with XIV, see the IBM Redpaper *XIV Security with Data-at-Rest Encryption*, REDP-5047.

- User auditing with access control lists and Idle timeout locking integration:

  The XIV Storage System Software offers the capability for user auditing with access control lists (ACLs) to provide more control and historical information.

  It also enforces UI locking after a period of idle time specified by administrator, so it requires re-authentication.

- Mobile push notifications

  Enables XIV clients to receive high severity events on mobile devices (iPhone, Android platforms).

- Support for IBM Tivoli® Storage Productivity Center:

  Tivoli Storage Productivity Center can discover IBM XIV systems and all internal components and manage capacity for storage pools, including allocated, deallocated, and available capacity with historical trending on use. It can also receive events and define policy-based alerts based on user-defined triggers and thresholds.

- Host Rate Limiting: Quality of service (QoS):

  The XIV Storage System system resources, such as storage and cache, constitute a virtualized environment that is shared by all hosts and applications. This approach lends itself exceptionally well to accommodate high performance requirements for multiple applications with similar performance objectives through fair resource allocation.

  XIV System Code release 11.5.0 introduced Domain based, and Pool based rate limiting.

  In environments with applications that have various performance objectives, the XIV QoS feature enables the client to restrict the system's processing power allocated to specified host applications, maximizing the power available for applications requiring the utmost performance.

- Support for IPv6 and IPSec:

  IPv6 and IPSec are supported in XIV Gen3 starting with software Version 11.1.*x*. The IPv6 support applies only to management ports. The internal XIV addresses are still IPv4 (no change).

  Starting with Version 11.1.1, the XIV Gen3 offers US Government (USG) IPv6 compliance.

- XIV is well-suited for cloud deployments:

  Both XIV Cloud Storage for Service Providers, or client hosted cloud solutions require elastic, rapid response, and dynamic demand that is captured within XIV's 489 TB linear scalability, and the immediate and automatic redistribution, without interruption, of added capacity. The grid design stripes data across all modules and spindles, incorporating data redundancy for consistent and predictable I/O performance that is always load balanced.

  XIV integration with the VMware vCloud Suite, by using the IBM Storage Integration Server, empowers cloud providers with simplicity, flexibility, and efficiency within the VMware cloud management suite including VASA, VCO, vCOPs, vCAC. In addition, XIV supports the VMware vStorage API Array Integration (VAAI) which moves the storage-related tasks that were previously performed by VMware hosts onto the storage system. Transferring the processing burden reduces performance impact; speeds processing; frees VMware for more mission-critical tasks, such as adding applications; simplifies management; and positions the virtual environment for additional capacity and scalability. When hardware acceleration is enabled with XIV System Storage, operations

such as VM provisioning, VM cloning, and VM migration complete dramatically faster, and with minimal impact to the ESX server, increasing scalability and storage performance. For more information about VMware topics, see the IBM Redpaper, *IBM XIV Storage System in a VMware Environment,* REDP-4965.

XIV supports IBM Hyper-Scale Mobility and IBM Hyper-Scale Consistency for scale-out needs. (See the IBM Redpaper, *IBM Hyper-Scale for the XIV Storage System*, REDP-5053).

XIV can automate storage discovery and provisioning in OpenStack environments using OpenStack Cinder integration. (See the IBM Redpaper, *Using the IBM XIV Storage System in OpenStack Cloud Environments*, REDP-4971.)

XIV can automate custom storage operation using the XIV RESTful API. (See the IBM Redpaper, *RESTful API Support in XIV*, REDP-5064).

1.6.1 IBM XIV Storage Management tools

The IBM XIV Management tools include the XIV Storage Management GUI, IBM XIV online monitoring tool (XIV Top), and IBM XIV Storage System Command-Line Interface (XCLI).

Along with IBM XIV Management Tools version 4.4, the IBM Hyper-Scale Manager v1.5.x was released. The Hyper-Scale Manager reduces operational complexity and enhances capacity planning through integrated management for large and multi-site XIV deployments.

The Hyper-Scale Manager can be run as a virtual appliance above an ESX server (VMware VSphere Hypervisor only) or as an application on a preinstalled Red Hat Enterprise Linux server.

For more information about the Hyper-Scale Manager, see the IBM Redpaper titled *IBM Hyper-Scale in XIV Storage, REDP-5053*.

The other XIV Management Tools (GUI, XIV Top, and XCLI) are available for various operating system platforms from the IBM System Storage Interoperation Center (SSIC):

`http://www.ibm.com/systems/support/storage/ssic/interoperability.wss`

The management tools, including the Hyper-Scale Manager and Host Attachment Kit (HAK), can be downloaded from IBM Fix Central:

`http://www.ibm.com/support/fixcentral/`

The XIV Storage Management GUI, XIV Top, and XCLI tools are bundled in a single package that can be downloaded for each supported operating system. There is a separate XCLI package for IBM AIX®, Linux, Solaris, and HPUX.

There is a mobile monitoring dashboard version available for iPhone, iPad, and Android devices.

IBM XIV Storage Management GUI

The IBM XIV Storage Management GUI acts as the management console for the storage system. A simple and intuitive GUI enables storage administrators to manage and monitor all system aspects easily, with almost no learning curve.

The XIV Storage Management GUI also contains a demonstration mode. To use the demonstration mode, after the initial GUI program launch, select **Demo** for the Mode and then click **Login**, as shown in Figure 1-2. A password is not required. There is also a Manager mode that you can use to activate the Hyper-Scale Storage Manager.

Storage Management
User:
xiv_development
Password:
Mode:
Direct
Manager (1): 9.155.59.232
Direct
Demo
Licensed Materials - Property of IBM Corporation and other(s). Copyright © 2008, 2012. IBM and XIV are registered trademarks of IBM Corporation in the United States, other countries, or both.

Figure 1-2 XIV GUI Demo Mode

Figure 1-3 shows one of the top-level configuration windows where you can also see flash (SSD) status.

Figure 1-3 The XIV Storage Management GUI

IBM Hyper-Scale Manager GUI

The IBM Hyper-Scale Manager GUI acts as the management console for multiple XIV systems. See Figure 1-4.

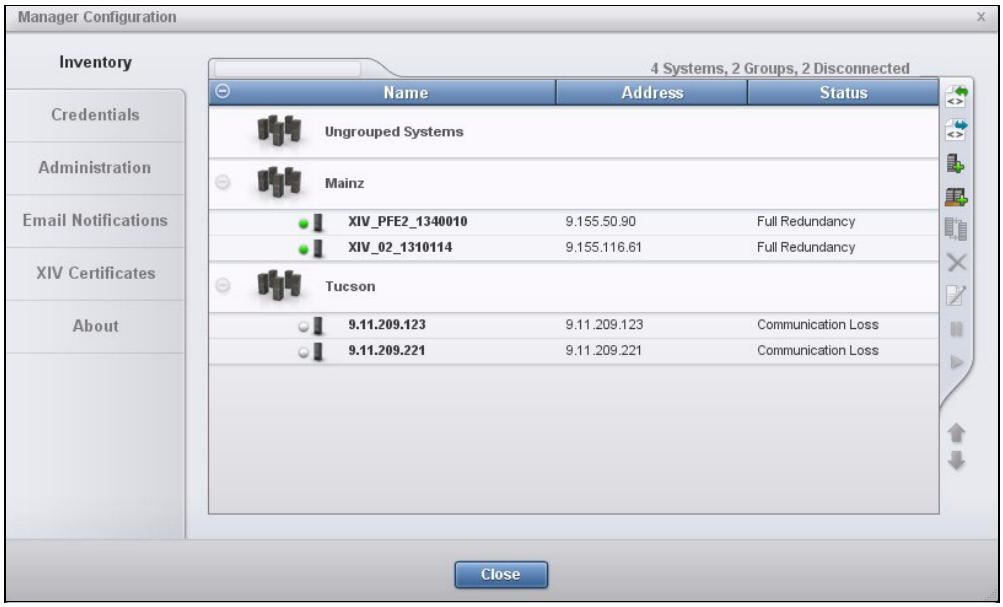


Figure 1-4 Hyper-Scale Manager GUI

XIV Storage Management online monitoring with XIV Top

With the XIV Top application, you can view and monitor performance information for defined volumes and hosts in real time. See Figure 1-5 for an illustration.

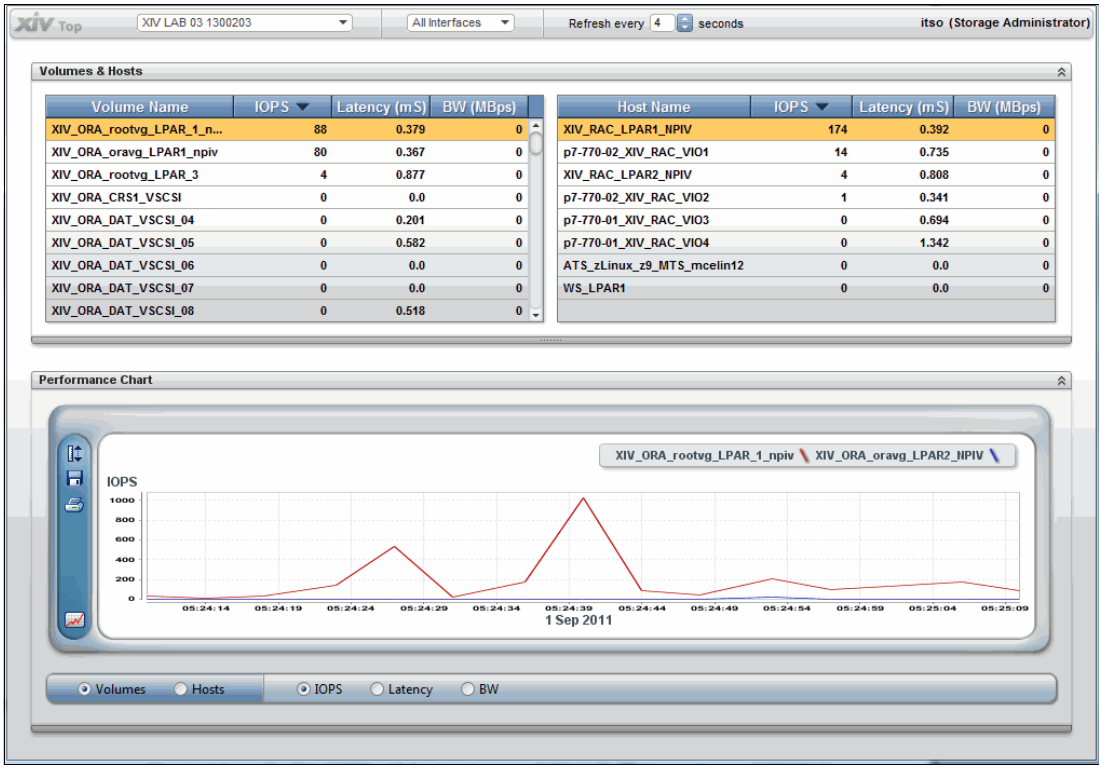


Figure 1-5 XIV Top

The tool quickly displays the volumes or hosts that are consuming the most system resources at any given time. Information can be sorted in various ways and saved to a file for use in spreadsheet applications.

XIV Storage Management XCLI

The XIV Storage Management XCLI is a comprehensive command-line interface to configure and monitor the system. All the functions available in the XIV Storage Management GUI are also available in the XCLI. The XCLI can be used in a shell environment to interactively configure the system or as part of a script to perform lengthy or complex tasks.

Example 1-1 shows an XCLI command being run in a Windows DOS shell.

Example 1-1 XCLI commands in a Windows DOS shell

```
C:\>xcli -c XIV_LAB config_get
Name                                 Value
dns_primary                          192.168.1.2
dns_secondary                        192.168.1.3
system_name                          XIV LAB 01 EBC
...
...
internal_email_subject_format        2810-A14: 6012345: {severity}: {description}
iscsi_name                           iqn.2005-10.com.xivstorage:000105
maximal_snapshot_deletion_priority   4
timezone                             -7200
fc_proof                             yes
ntp_server                           192.168.1.100
ups_control                          yes
support_center_port_type             Management
```

The XCLI is installed as part of the FULL GUI installation, or it can be installed alone.

1.7 Host support

The XIV Storage System can be attached to various host operating systems, including the following ones:

- Microsoft Windows
- VMware ESX
- IBM AIX
- HPUX
- Linux (also zLinux and Linux on POWER®)
- Oracle Solaris

For details about each operating system and versions supported, see the IBM XIV interoperability matrix or the IBM System Storage Interoperation Center (SSIC) at the following website:

http://www.ibm.com/systems/support/storage/config/ssic/index.jsp

There are also various software *agents* available to help you manage environments that connect to an XIV Storage System.

An IBM storage driver for OpenStack is also available. For more information about this topic, see the IBM Redpaper publication, *Using the IBM XIV Storage System in OpenStack Cloud Environments*, REDP-4971.

You can also consult the following resources:

- IBM Storage Integration Server

  http://ibm.co/1wgeriO

  This is a free-of-charge,centralized server that consolidates IBM storage provisioning, automation, and monitoring through a unified server platform. It is compatible with Red Had Enterprise Linux (RHEL) on a local or virtual host.

  Version 1.5 incorporates VMware APIs for Storage Awareness (VASA), VMware vSphere Web Client, vCenter Orchestrator (vCO), vCOPs, and VMware vCloud Automation Center (vCAC).

- IBM XIV Management Console for VMware vCenter

  It integrates XIV storage with the VSphere console for more efficient management of VMWare components on XIV storage. (They link through the IBM Support Portal.)

- VMware vCenter Operations Manager (vCOPs)

  https://my.vmware.com/web/vmware/downloads

- VMware vCenter Site Recovery Manager Storage Replication Adapters

  https://my.vmware.com/web/vmware/downloads

  Storage Replication Adapters are software modules for VMware Site Recovery Manager that simplify the use of storage replication software with VMware vSphere.

- Fix Central provides fixes and updates for your XIV System software, hardware, and operating system at this website:

  http://www.ibm.com/support/fixcentral/swg/selectFixes?parent=Enterprise+Storage+Servers&product=ibm/Storage_Disk/XIV+Storage+System+%282810,+2812%29&release=All&platform=All&function=all

- IBM Storage enabler for Windows Failover Clustering Agent

  The IBM Storage enabler for Windows Failover Clustering Agent maintains volume mirroring between two IBM Storage systems and Microsoft failover clusters.

- IBM XIV Provider for Windows Volume Shadow Copy Service (VSS)

  This integrates XIV Storage System snapshot capabilities with Microsoft Windows Volume Shadow Copy Services. (They link through the IBM Support Portal.)

- IBM Storage Management Pack for Microsoft Systems Center Operations Manager

  The Storage Management Pack for Microsoft Systems Center Operations Manager (SCOM) is a set of software modules, or management packs, thatyou can use to access and monitor IBM storage systemsby using the host-based Microsoft SCOM interface.

- Tivoli Storage Productivity Center Storage Resource agents

  http://ibm.co/10eRz8F

  These agents perform the functions of Data agents and Fabric agents. They support XIV Storage System provisioning, Data Path Explorer, and performance management reports.

2

IBM XIV Storage System logical architecture and concepts

In this chapter, we elaborate on various underlying design and architectural concepts of the IBM XIV Storage System that was introduced in Chapter 1, “IBM XIV Gen3 Storage System overview” on page 1.

This chapter covers the following topics:

- 2.1, “Architecture overview” on page 18
- 2.2, “Parallelism” on page 21
- 2.3, “Full storage virtualization” on page 24
- 2.7, “Capacity allocation and thin provisioning” on page 37
- 2.8, “Flash caching architecture” on page 43
- 2.9, “Encryption for data at rest” on page 45
- 2.10, “Reliability, availability, and serviceability (RAS)” on page 46

© Copyright IBM Corp. 2014. All rights reserved.

2.1 Architecture overview

IBM XIV Storage System Gen3 carries on the XIV tradition of providing several architecture design factors that contribute to its unique operational capabilities:

- Virtual, grid-based array design (to minimize storage complexity)
- Self-tuning and self-healing (to deliver performance and availability without manual tuning)
- Easy storage management and provisioning (an intuitive graphical user interface (GUI) to simplify the storage operation within a virtualized storage architecture)
- Enhanced connectivity and reliability (to strengthen the end-to-end performance without disruption to the service)
- Innovative caching with optional flash cache, coupled with high-density storage (to reduce the footprint and address energy efficiency while sustaining performance)

2.1.1 Features

The XIV Storage System architecture incorporates various features designed to uniformly distribute data across internal resources. This unique data distribution method fundamentally differentiates the XIV Storage System from conventional storage subsystems, offering numerous availability, performance, and management benefits across both physical and logical elements of the system.

The XIV Storage System has a native virtual storage design that is both efficient and simple to use. It eliminates all physical disk Redundant Array of Independent Disks (RAID) management tasks. The XIV is an enterprise-class storage system in which storage management involves defining capacity (a logical unit number, or LUN, or a volume) and assigning the capacity to a host system.

The XIV Storage System hardware architecture is a dramatic shift from traditional RAID-based parity data protection. From a performance perspective, the XIV Storage System can *automatically* involve all of the disk drives, the whole system cache including its optional flash cache extension, and the processors in servicing I/O operations.

Scalability is also a strength of the XIV Storage System. The XIV Storage System is able to scale without administrative involvement to redistribute data or tune for optimum performance. It takes care of these tasks automatically.

2.1.2 Hardware elements

To convey the concepts that the XIV Storage System architecture is based on, it is useful to first look at the physical design.

The XIV configuration includes data modules, interface modules, interconnect switches, and uninterruptible power supply (UPS) units. For details and components of the physical system architecture, see Chapter 3, “IBM XIV architecture, components, and planning” on page 61.

Modules

The primary components of the XIV Storage System are known as *modules*. Modules provide processing, cache (including optional flash cache), and host interfaces. They are composed of “off the shelf” systems that are based on Intel technology.

The modules are redundantly connected to one another through an internal switched network, as shown in Figure 2-1.

All of the modules work together, concurrently, as elements of a grid architecture. This helps the system harness the parallelism that is inherent in such a distributed computing environment.

We describe the grid architecture in 2.2, “Parallelism” on page 21.

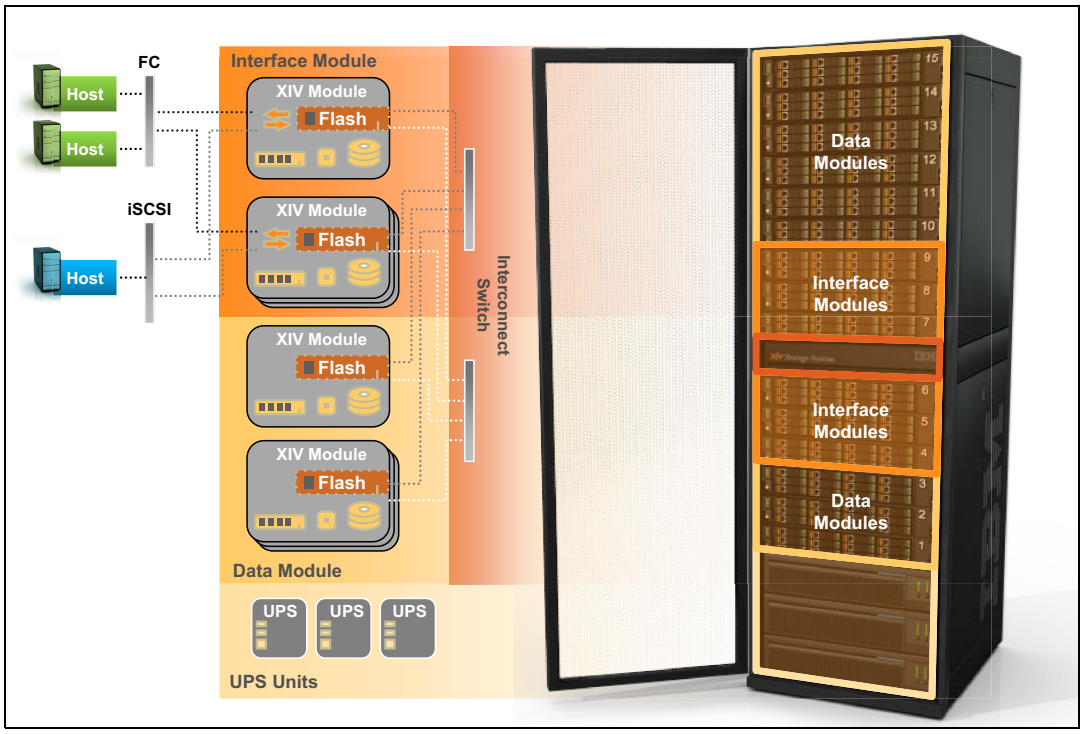


Figure 2-1 XIV Storage System major hardware elements

Important: Flash cache is an optional feature. However, if you opt for the flash cache extension, flash cache devices must be installed in *all* modules.

Data modules

At a conceptual level, data modules function as the elementary building blocks of the system. They provide storage capacity, processing power, and caching functions in addition to advanced system-managed services. The data modules' abilities to share and manage system software and services are key elements of the physical architecture, as shown in Figure 2-2.

Note: Figure 2-2 depicts the conceptual architecture only. Do not misinterpret the number of connections or modules and other representations as a precise hardware layout.

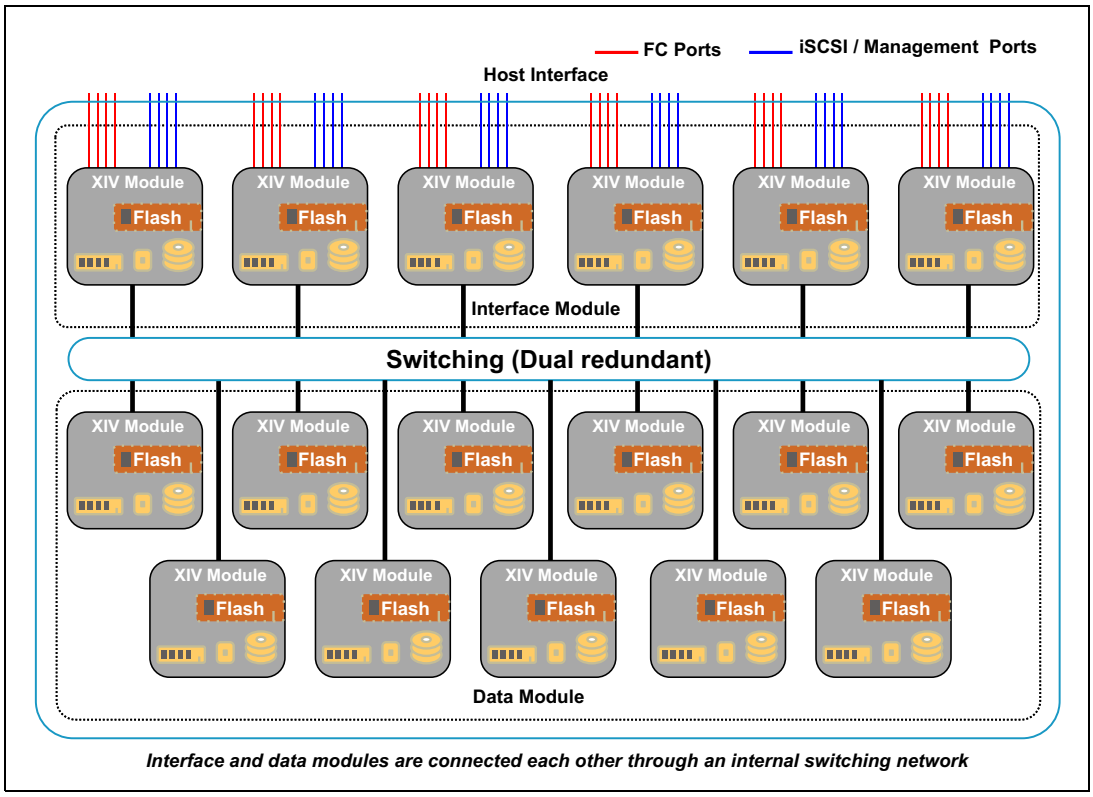


Figure 2-2 XIV Storage System architectural overview

Interface modules

Interface modules are equivalent to data modules in all aspects, with the following exceptions:

- In addition to disk, cache, and processing resources, interface modules include both Fibre Channel and IP network Small Computer System Interface (iSCSI) interfaces for host system connectivity, remote mirroring, and data migration activities. Figure 2-2 conceptually illustrates the placement of interface modules within the topology of the XIV Storage System architecture.
- The system services and software functions associated with managing external I/O is located exclusively on the interface modules.

Module interconnect switches

The XIV Storage System contains a redundant switched network, based on InfiniBand, that transmits both data and metadata traffic between the modules. Traffic can flow between two interface modules, between two data modules, and between an interface module and a data module.

2.2 Parallelism

The concept of *parallelism* pervades all aspects of the XIV Storage System architecture by using a balanced, redundant data distribution scheme with a pool of distributed (or grid) computing resources. To explain the principle of parallelism further, it is helpful to consider the ramifications of both the hardware and software implementations independently. We then examine virtualization principles in 2.3, “Full storage virtualization” on page 24.

Important: The XIV Storage System uses parallelism at *both* hardware and software levels.

2.2.1 Hardware parallelism and grid architecture

The XIV grid design (Figure 2-3 on page 22) has the following characteristics:

- Both interface and data modules work together in a distributed computing sense. However, the interface modules also have additional functions and features associated with host system and remote mirroring connectivity.
- The modules communicate with each other through an internal, redundant switched network.
- The software services and distributed computing algorithms running within the modules collectively manage all aspects of the operating environment.

Design principles

The XIV Storage System grid architecture, by virtue of its distributed topology, ensures that the following design principles are possible:

- Performance:
  - The relative effect of the loss of a single component is minimized.
  - All modules are able to participate equally in servicing the total workload.

    This design principle is true regardless of access patterns. The system architecture inherently enables excellent load balancing, even if certain applications access certain volumes, or certain parts within a volume, more frequently.
- Compatibility:
  - Modules consist of standard “off the shelf” components.

    Because components are not specifically engineered for the system, the resources and time required for the development of newer hardware technologies are minimized. This benefit, coupled with the efficient integration of computing resources into the grid architecture, enables the system to realize the rapid adoption of newer hardware technologies available without the need to deploy a whole new subsystem.
- Scalability:
  - Computing resources can be dynamically changed.
  - The architecture can be “scaled out” by adding new modules to accommodate both new capacity and new performance demands.
  - The architecture can also be “scaled up” by adding more modules, up to a total of 15 modules.

Figure 2-3 contains a conceptual depiction of the scalable grid architecture of the XIV.

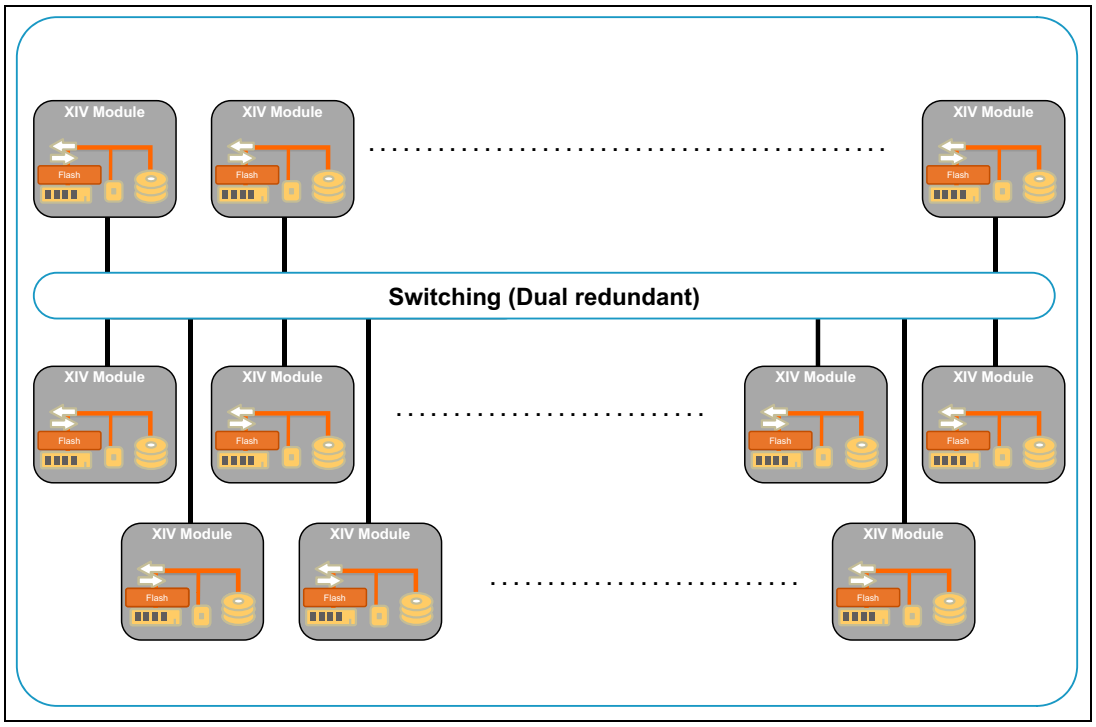


Figure 2-3 XIV Storage System scalable conceptual grid architecture

Proportional scalability

Within the XIV Storage System, each module contains all of the pertinent hardware elements that are necessary for a grid topology (processing, caching, and storage). All modules are connected through a scalable network. This aspect of the grid infrastructure enables the relative proportions of cache, processor, disk, and interconnect bandwidth to remain optimal even when modules are added or removed:

- Linear cache growth: The total system cache size and cache bandwidth increase linearly with disk capacity because every module is a self-contained computing resource that houses its own cache. The cache bandwidth scales linearly in terms of both host-to-cache and cache-to-disk throughput, and the close proximity of cache, processor, and disk is maintained.

  Cache: Flash cache also increases linearly with total memory capacity. Flash cache drives are not used to expand the disk capacity, but rather to extend the memory capacity allocated to the caching of read I/Os.

- Proportional interface growth: Interface modules house Ethernet and Fibre Channel host interfaces and are able to access not only the local resources within the module, but also the entire system. With every interface module added, the system proportionally scales both the number of host interfaces and the bandwidth to the internal resources.
- Constant switching capacity: The internal switching capacity scales proportionally as the system grows, preventing bottlenecks regardless of the number of modules. This capability ensures that internal throughput scales proportionally to capacity.

- Embedded processing power: Because each module incorporates its own processing power with cache and disk components, the ability of the system to perform processor-intensive tasks, such as aggressive prefetch caching, sophisticated cache updates, snapshot management, and data distribution, is always maintained regardless of the system capacity.

2.2.2 Software parallelism

In addition to the hardware parallelism, the XIV Storage System also employs sophisticated algorithms to achieve optimal software parallelism.

Modular software design

The XIV Storage System internal operating environment consists of a set of software functions that are loosely coupled with the hardware modules. These software functions reside on one or more modules and can be redistributed among modules as required, therefore ensuring resiliency under changing hardware conditions.

An example of this modular design is located specifically in the interface modules. All six interface modules actively manage system services and software functions associated with managing external I/O. Also, three of the interface modules deliver the system's management interface service for use with the XIV Storage System.

Data distribution algorithms

Data is distributed across all drives in a *pseudo-random* fashion. Patented algorithms provide a uniform yet random distribution of data, which is divided into 1 MB partitions across all available disks, to maintain data resilience and redundancy.

Figure 2-4 shows how all drives are used evenly by partition units regardless of applications or the size of assigned logical volumes. The storage administrator does not need to worry about data placement impact on performance.

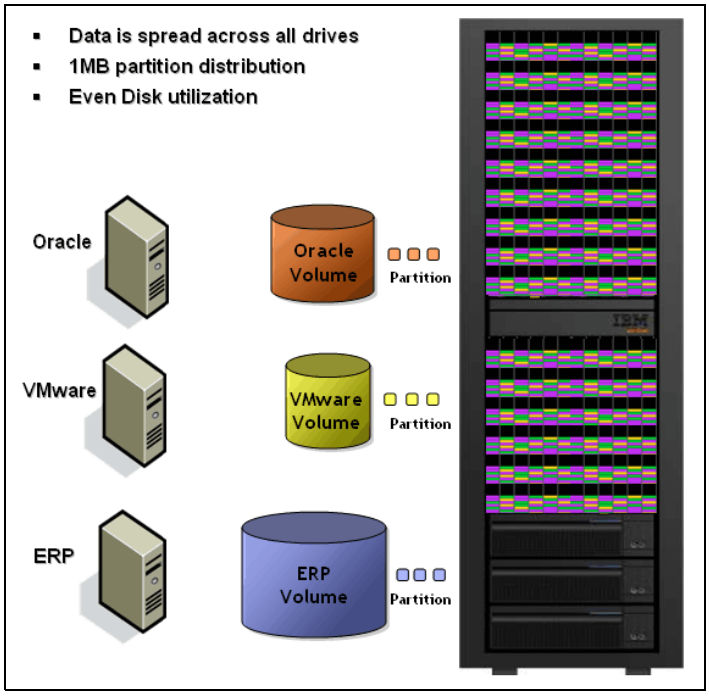


Figure 2-4 XIV Storage System pseudo-random data distribution

For more details about the topic of data distribution and storage virtualization, see 2.4, “Logical system concepts” on page 26.

2.3 Full storage virtualization

The data distribution algorithms employed by the XIV Storage System are innovative and unique in that they are deeply integrated into the system architecture itself, instead of at the host or storage area network level.

To appreciate the value inherent to the virtualization design that is used by the XIV Storage System, consider the various aspects of the physical and logical relationships that make up conventional storage subsystems. Specifically, traditional subsystems rely on storage administrators to plan the relationship between logical structures, such as arrays and volumes, and physical resources, such as disk packs and drives, to strategically balance workloads, meet capacity demands, eliminate hotspots, and provide adequate performance.

2.3.1 XIV Storage System virtualization design

The implementation of full storage virtualization employed by the XIV Storage System eliminates many of the potential operational drawbacks that can be present with conventional storage subsystems, while maximizing the overall usefulness of the subsystem.

The XIV Storage System virtualization offers the following benefits:

- Easier volume management:
  - Logical volume placement is driven by the distribution algorithms, freeing the storage administrator from planning and maintaining volume layout. The data distribution algorithms manage all of the data in the system collectively without deference to specific logical volume definitions.
  - Any interaction, whether host or system driven, with a specific logical volume in the system is inherently handled by all resources. It harnesses all storage capacity, all internal bandwidth, and all processing power currently available in the system.
  - Logical volumes are not exclusively associated with a subset of physical resources:
    - Logical volumes can be dynamically resized.
    - Logical volumes can be thinly provisioned, as described in 2.7, “Capacity allocation and thin provisioning” on page 37.
- Consistent performance and scalability:
  - Hardware resources are always used equally because all logical volumes always span all physical resources and are therefore able to reap the performance potential of the full system and maintain data integrity.
    - Virtualization algorithms automatically redistribute the logical volumes’ data and workload when new hardware is added, maintaining the system balance while preserving transparency to the attached hosts.
    - If there is a hardware failure, data is automatically, efficiently, and rapidly rebuilt across all the drives and modules in the system. This action preserves host transparency, equilibrium, and data redundancy at all times while virtually eliminating any performance penalty associated with traditional RAID rebuilds. This rebuild action works together with the same redistribution activity as when hardware is added.

  - There are no “pockets” of capacity, “orphaned” disk space, or resources that are inaccessible because of array mapping constraints or data placement.

- Flexible snapshots:
  - Full storage virtualization incorporates snapshots that are differential in nature. Only updated data uses physical capacity, which makes the following capabilities possible:
    - Multiple concurrent snapshots are possible because a snapshot uses physical space only after a change has occurred on the source.
    - Multiple snapshots of a single master volume can exist independently of each other.
    - Snapshots can be cascaded, in effect creating snapshots of snapshots.
  - Creation and deletion of snapshots do not require data to be copied and occur immediately.
  - When updates occur to master volumes, the system’s virtualized logical structure enables it to preserve the original point-in-time data associated with any dependent snapshots by redirecting the update to a new physical location on disk. This process, referred to as *redirect on write,* occurs transparently from the host perspective and uses the virtualized remapping of the updated data to minimize any performance impact associated with preserving snapshots, regardless of the number of snapshots defined for a specified master volume.

    Snapshots: The XIV snapshot process uses *redirect on write*, which is more efficient than the *copy on write* that is used by many other storage subsystems.

- Data migration efficiency:
  - XIV supports thin provisioning. When migrating from a system that supports only regular (or thick) provisioning, XIV allows thick-to-thin provisioning of capacity. Thin-provisioned capacity is described in 2.7, “Capacity allocation and thin provisioning” on page 37.
  - Because of the XIV pseudo-random distribution of data, the performance impact of data migration on production activity is minimized because the load is spread evenly over all resources.
- Automatic and dynamic flash cache use for improved performance:
  - XIV supports one flash cache extension per module. After the optional flash cache device is installed (it must be present in all modules), the XIV caching algorithm manages the flash cache as an extended, read-only cache to provide a performance boost for specific workloads.
  - No need for any specific consideration for flash cache use:
    - No software
    - No decision
    - No tuning
  - XIV allocates flash read-cache space dynamically according to detected workload patterns.
  - XIV supports optional, selective flash cache activation per logical volume (host volume).

– XIV supports nondisruptive online upgrade.

> **Flash caching:** For details about flash caching, see 2.8, “Flash caching architecture” on page 43, or consult the Redpaper publication, *Solid-State Drive Caching Implementation in the IBM XIV Storage System*, REDP-4842.

- Quality of service (QoS):

  The XIV disks and cache are shared at all times among all applications that are running on the various host servers attached to the XIV system. This approach greatly simplifies the tasks of the storage administrator. However, when the system is running at full capacity, it can have the adverse effect that noncritical applications are served at the same level as the business critical applications. The QoS feature allows you to prevent such situations. With QoS, the storage administrator can specify and enforce limits on the performance of low-priority workloads by limiting input/output operations per second (IOPS) and bandwidth on a per-host basis. The multitenancy feature introduced in XIV System software Version 11.5 extends QoS functions to the enforcement of limits upon hosts or storage pools and domains. For more information, see 2.4.3, “Multitenancy” on page 30 and to 4.7, “QoS feature” on page 170.

2.4 Logical system concepts

In this section, we elaborate on the logical system concepts, which form the basis for the system full storage virtualization.

2.4.1 Logical constructs

The XIV Storage System logical architecture incorporates constructs that underlie the storage virtualization and distribution of data, which are integral to its design. The logical structure of the system ensures that there is optimum granularity in the mapping of logical elements to both modules and individual physical disks, guaranteeing an equal distribution of data across all physical resources.

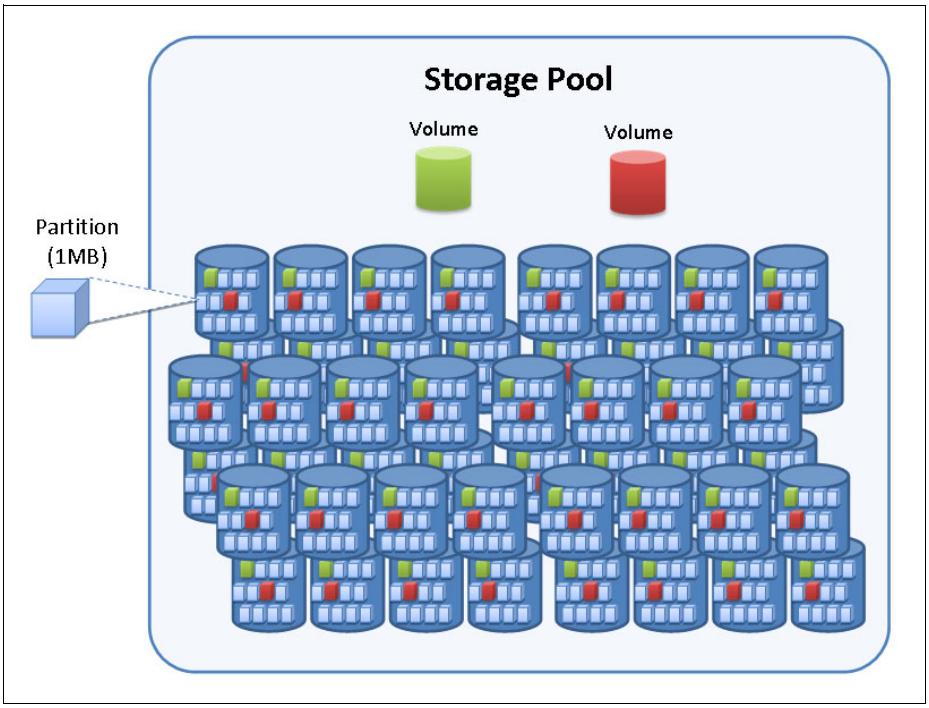


Figure 2-5 Logical system virtualization concept

Partitions

The fundamental building block of a logical volume is known as a *partition*. Partitions have the following characteristics on the XIV Storage System:

- All partitions are 1 MB (1024 KB) in size.
- A partition contains either a primary copy or secondary copy of data:
  - Each partition is mapped to a single physical disk:
    - This mapping is dynamically managed by the system through innovative data distribution algorithms to preserve data redundancy and equilibrium. For more information about the topic of data distribution, see “Logical volume layout on physical disks” on page 28.
    - The storage administrator has no control or knowledge of the specific mapping of partitions to drives.
  - Secondary copy partitions are always placed in a separate *module* from the one containing the primary copy partition.

Important: In the context of the XIV Storage System logical architecture, a partition consists of 1 MB (1024 KB) of data. Do not confuse this definition with other definitions of the term *partition*.

The diagram in Figure 2-4 on page 23 illustrates that data is uniformly yet randomly distributed over all disks. Each 1 MB of data is duplicated in a primary and secondary partition for redundancy. The system ensures that the primary partition and its corresponding secondary partition are never located within the same module to protect against a single point of failure.

Logical volumes

The XIV Storage System presents logical volumes to hosts in the same manner as conventional subsystems. However, both the granularity of logical volumes and the mapping of logical volumes to physical disks differ:

- Every logical volume consists of 1 MB (1024 KB) constructs of data known as partitions.
- The physical capacity associated with a logical volume is *always a multiple of 17 GB (decimal)* because the volume spans all physical drives in the system because of the grid nature of the architecture.

  Although it is possible to present a block-designated logical volume to a host that is *not* a multiple of 17 GB, the *actual physical space* that is allocated for the volume is always the sum of the minimum number of 17 GB increments needed to meet the block-designated capacity.

  Capacity: The initial physical capacity allocated by the system upon volume creation can be less than this amount, as described in “Logical and actual volume sizes” on page 38.

- A maximum total of 12,000 volumes (including snapshots) can be concurrently created on the system.

Storage pools

Storage pools are administrative boundaries that enable storage administrators to manage relationships between volumes and snapshots and to define separate capacity provisioning and snapshot requirements for separate applications or departments. Storage pools are not tied in any way to physical resources, nor are they part of the data distribution scheme. We describe storage pools and their associated concepts in 2.6, “Storage pool concepts” on page 35.

Snapshots

A *snapshot* represents a point-in-time copy of a volume. Snapshots are like volumes, except snapshots incorporate dependent relationships with their source volumes, which can be either logical volumes or other snapshots. Because they are not independent entities, a particular snapshot does not necessarily wholly consist of partitions that are unique to that snapshot. Conversely, a snapshot image does not share all of its partitions with its source volume if updates to the source occur after the snapshot was created.

2.4.2 Logical volume layout on physical disks

The XIV Storage System manages the distribution of logical volumes over physical disks and modules with a dynamic relationship between primary data partitions, secondary data partitions, and physical disks. This virtualization of resources in the XIV Storage System is governed by the data distribution algorithms.

Distribution table

The distribution table is created at system startup. It contains a map of every primary and secondary partition and the modules and physical disks where they are located. When hardware changes occur, a new distribution table is created and delivered to every module. Each module retains redundant copies of the distribution table.

Volume layout

At a conceptual level, the data distribution scheme can be thought of as a mixture of mirroring and striping. Although it is tempting to think of this scheme in the context of RAID 1+0 (10) or 0+1, the low-level virtualization implementation precludes the use of traditional RAID algorithms in the architecture.

As described in "Partitions" on page 27, the XIV Storage System architecture divides logical volumes into 1 MB partitions. This granularity and the mapping strategy are integral elements of the logical design that enable the system to realize the following features and benefits:

- Partitions that make up a volume are distributed on all disks by using a *pseudo-random distribution function*, which is described in 2.2.2, "Software parallelism" on page 23:
  - The distribution algorithms seek to preserve the equality of access among all physical disks under all conceivable conditions and volume access patterns. Essentially, although not truly random in nature, the distribution algorithms in combination with the system architecture preclude the occurrence of *hotspots*:
    - A fully configured XIV Storage System contains 180 disks, and each volume is allocated across at least 17 GB (decimal) of capacity that is distributed evenly across all disks.
    - Each logically adjacent partition on a volume is distributed across a separate disk. Partitions are *not* combined into groups before they are spread across the disks.
    - The pseudo-random distribution ensures that logically adjacent partitions are never striped sequentially across physically adjacent disks. For more about the partition mapping topology, see 2.2.2, "Software parallelism" on page 23.
  - Each disk has its data mirrored across all other disks, *excluding the disks in the same module*.
  - Each disk holds approximately one percent of any other disk in other modules.
  - Disks have an equal probability of being accessed regardless of aggregate workload access patterns.
- The following information is described in "XIV Storage System virtualization design" on page 24:
  - The storage system administrator does *not* plan the layout of volumes on the modules.
  - If there is space available, volumes can always be added or resized instantly with a negligible impact on performance.
  - There are no unusable pockets of capacity known as *orphaned spaces*.
- When the system is scaled out through the addition of modules, a new data distribution is created, where just a minimum number of partitions are moved to the newly allocated capacity to arrive at the new distribution table.

  The new capacity is fully used within a few hours and with no need for any administrative intervention. Therefore, the system automatically returns to a state of equilibrium among all resources.
- If a drive or module fails or is phased out, a new XIV Storage System data distribution is created where data in non-redundant partitions is copied and redistributed across the remaining modules and drives.

  The system rapidly returns to a state in which all partitions are again redundant because all disks and modules participate in re-creating the necessary partitions.

2.4.3 Multitenancy

Today, many organizations are moving toward cloud-based infrastructures, where both compute and storage resources are viewed as commodity services. Users of these resources expect a pay-as-you-go model, where the resources that they require are rapidly available when needed, but they pay only for the resources that they require, nothing more. For storage providers, this shift represents a significant challenge. From both capital expenditure and resource management (personnel) perspectives, storage consolidation becomes an important consideration. This new approach necessitates storage systems that can easily and securely support the provisioning, administration, and service level management of storage resources for multiple resource to be used concurrently.

XIV multitenancy, introduced in XIV Storage Software v11.5, brings flexibility and simplicity to management of tenant data and storage resources across multiple XIV systems through these methods:

- Secure division and isolation of XIV storage resources among numerous tenants
- Simple, quick delegation of administration tasks and role-based permissions
- Simple, rapid deployment without the need for extensive planning and tuning and the ability to upgrade in the field

Domains

XIV multitenancy is based upon the concept of *domains*, where an XIV Storage System is logically partitioned into one or more independent containers, each with its own assigned administrators. This enables secure isolation form other domains of the logical entities contained within a domain. A user who is associated with a single domain has no knowledge of the other domains that exist on the system nor about the pools or volumes associated with those domains. Domains can be associated with these entities:

- Users and user groups
- Storage pools (and, inherently, the volumes that they contain)
- Hosts and clusters
- Remote mirror targets

Note: Although a storage pool (and the volumes it contains) can be associated with only a *single* domain, users (and user groups), hosts (and host clusters), and remote mirror targets can be associated with *multiple* domains.

From a conceptual perspective, an end user's interactions (for example, those of a storage administrator) with the XIV Storage System can be viewed as *actions* performed upon *objects*. Examples of these object/action pairs might include:

- Creating a storage pool
- Resizing a volume
- Mapping a volume to a host
- Viewing the properties of a pool or volume

The XIV Storage System uses a role-based access control (RBAC) model to control *what actions* a specific system user can perform. The predefined roles available in the system are discussed in more detail in “User roles” on page 221 and in “LDAP user roles” on page 266.

Multitenancy, through the use of domains, introduces the ability to control *what objects* a specific system user can perform those actions upon.

Note: When considering multitenancy, do not confuse the use of the term *domain* with a fully-qualified domain name (FQDN) or directory services domain. In the context of multitenancy, a domain is simply a logical construct that allows partitioning (logically, not physically) of XIV Storage System resources.

It is important to understand that the implementation of this domain construct within the XIV Storage System occurs at the management level. All physical system resources (disks, CPUs, memory, and interfaces) are shared among domains. As such, domain administrators do not have the ability to modify physical system resources. However, they can be notified of events relating to physical system attributes, because these events might affect the objects within their domain. For example, a domain administrator can be alerted upon the failure of a drive or the disconnection of an interface link.

Domain creation

The creation of domains within the XIV Storage System is performed by a *global storage administrator*, which is a user with the role of storage administrator who is not associated with any domains. By default, the built-in *admin* user account is a global storage administrator.

When creating a domain, the global administrator assigns system resources to that domain. There are certain resources within the XIV Storage System that have finite limits. Examples include storage capacity and number of volumes. As such, upon creation of a domain, the global administrator must determine the quantity of these finite resources to assign to the domain. For more information about domain creation and the assignment of system resources to a domain, see 4.8, “Multitenancy” on page 174.

Domain attributes

Domains includes these important characteristics:

- A domain is a logical partition of the system’s resources. It represents a subset of the system’s resources. These resources include, but are not limited to, storage pools, hosts, mirror targets.
- Users can be assigned to zero or more domains. A user assigned to zero domains is considered a global user and has access to all system resources that are not associated exclusively with a domain.
- A domain restricts the resources that a user can manage. A user can manage only the parts of the system that are associated with the domains that he or she is associated with, as depicted in Figure 2-6 on page 32.

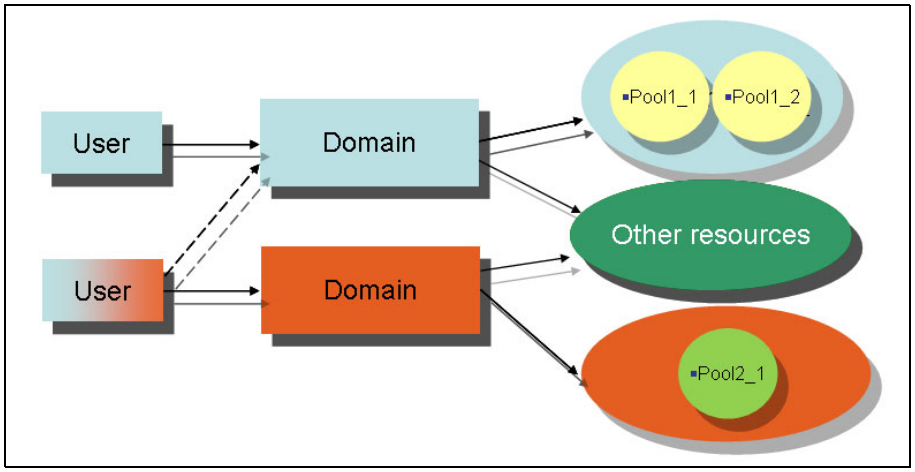


Figure 2-6 User view of domains and resources

- A *domain administrator* is a user who is associated with a domain or domains. The domain administrator is restricted to performing operations on objects associated with a specific domain or domains that he or she is associated with.

  A domain administrator can be assigned to manage multiple domains, and a domain can be assigned to be managed by multiple administrators.
- There is no visibility between domains. Domain administrators are not informed of resources outside of their domains. These resources and their related events or alerts are not displayed on lists.
- Within a domain, the storage, security, and application administrators, as well as read-only users retain their rights to perform the same operations that they have in an environment that is not based on domains. However, the operations are limited to the resources within the domain.
- Storage allocated to a domain can be used to create pools within that domain.
- A pool (and its volumes) can be associated with only one domain.
- Hosts, clusters, and targets can be associated with domains in a non-exclusive (shared) manner.
- Resources that are not associated with any domain are accessible only to global administrators.
- Quality of service can be set independently for each domain and can be used to limit the bandwidth or IOPS allowed for a domain or for the storage pools within a domain.

2.5 System capacity

The XIV Storage System has two different concepts of storage capacity:

- Net usable capacity
- System reserve capacity

2.5.1 Net usable capacity

Net usable capacity is the available data storage capacity that comes with the purchased XIV System configuration. It is a function of the drive size (1 TB, 2 TB, 3 TB, or 4 TB) and the number of modules.

The calculation of the net usable capacity of the system consists of the total disk count, less disk space reserved for sparing (which is the equivalent of one module plus three more disks). This number is then multiplied by the amount of capacity on each disk drive that is dedicated to data (98%), and reduced by a factor of 50% to account for data mirroring achieved by the secondary copy of data.

For example, an XIV Storage System with 15 modules populated with 2 TB disk drives has a net usable capacity of approximately 161 TB based on the following formula:

(((180 - 15) * 2 TB) * .98) / 2 = 161.7

The net usable capacity depends on the module configuration and disk drive type. For more details, see Table 2-1.

Table 2-1 Net usable capacity depends on module configuration and disk drive type (TB, decimal)

Total number of modules	**6**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Net capacity with 1 TB disk drives	28 TB	44 TB	51 TB	56 TB	63 TB	67 TB	75 TB	80 TB
Net capacity with 2 TB disk drives	55 TB	88 TB	102 TB	111 TB	125 TB	134 TB	149 TB	161 TB
Net capacity with 3 TB disk drives	84 TB	132 TB	154 TB	168 TB	190 TB	203 TB	225 TB	243 TB
Net capacity with 4 TB disk drives	112 TB	177 TB	207 TB	225 TB	254 TB	272 TB	301 TB	325 TB
Net capacity with 6 TB disk drives	169 TB	267 TB	311 TB	338 TB	382 TB	409 TB	453 TB	489 TB

Note: The system defines capacity by using decimal metrics. One GB is 1 000 000 000 bytes using decimal metrics. By contrast, 1 GiB is 1 073 741 824 bytes using binary metrics.

The XIV Storage System reserves physical disk capacity for:

- Global spare capacity
- Metadata, including statistics and traces
- Mirrored copies of data

2.5.2 Global spare capacity

The dynamically balanced distribution of data across all physical resources by definition eliminates the need of dedicated spare drives that are necessary with conventional RAID technologies. Instead, the XIV Storage System reserves capacity on each disk drive to provide adequate space for the redistribution or rebuilding of redundant data in the event of a hardware failure.

This global spare capacity approach offers advantages over dedicated hot spare drives, which are used only upon failure and are not used otherwise, therefore reducing the number of spindles that the system can use for better performance. Also, those non-operating disks are not typically subjected to background scrubbing processes. In XIV, all disk drives are operating all the time and are subject to examination, which helps detect potential reliability issues with drives.

The global reserved space includes sufficient capacity to withstand the failure of a full module plus three additional disk drives, and still allow the system to create the necessary partition copies to return to full redundancy.

Tolerance of failures: A fully used system tolerates multiple hardware failures, including up to an entire module with three subsequent drive failures *outside of the failed module*. If the system is less than 100% full, it can sustain more subsequent failures based on the amount of unused disk space that is allocated in the event of failure as a spare capacity. For a thorough description of how the system uses and manages reserve capacity under specific hardware failure scenarios, see 2.10.2, “Preserving data redundancy: Rebuilding and redistributing” on page 51.

Snapshots: The IBM XIV Storage System does not manage a separate global reserved space for snapshots. For more about this, see 2.6.3, “Storage pool relationships and rules” on page 36.

2.5.3 Metadata and system reserve

The system reserves roughly 4% of the physical capacity for statistics and traces, and the distribution table.

2.5.4 Mirrored copies of data

The mirrored copies of data are the secondary partitions described in “Partitions” on page 27.

2.6 Storage pool concepts

The concept of *storage pools* is purely administrative. Although the hardware resources within the XIV Storage System are virtualized in a global sense, the available capacity in the system can be administratively portioned into separate and independent storage pools. Essentially, storage pools function as a way to effectively manage a related group of similarly provisioned logical volumes and their snapshots.

2.6.1 Improved management of storage space

Storage pools form the basis for controlling the use of storage space by imposing a capacity quota on specific applications, a group of applications, or departments. This enables isolated management of relationships within the associated group of logical volumes and snapshots. Storage pools can also be uniquely associated with a domain.

A *logical volume* is defined within the context of only one storage pool. Because storage pools are logical constructs, a volume and any snapshots associated with it can be moved to any other storage pool (within the same domain) if there is sufficient space within the target storage pool.

As a benefit of the system virtualization, there are no limitations on the associations between logical volumes and storage pools. In fact, manipulation of storage pools consists exclusively of metadata transactions and does not trigger any copying of data. Therefore, changes are completed instantly and without any system performance degradation.

2.6.2 Consistency groups

A *consistency group* is a group of volumes of which a snapshot can be made at the same point in time, therefore ensuring a consistent image of all volumes within the group at that time. The concept of a consistency group is common among storage subsystems in which it is necessary to perform concurrent operations collectively across a set of volumes so that the result of the operation preserves the consistency among volumes. For example, effective storage management activities for applications that span multiple volumes, or creating point-in-time backups, is not possible without first employing consistency groups.

This consistency between the volumes in the group is paramount to maintaining data integrity from the application perspective. By first grouping the application volumes into a consistency group, it is possible to later capture a consistent state of all volumes within that group at a specified point-in-time using a special snapshot command for consistency groups.

Issuing this type of a command results in the following process:

1. Complete and destage writes across the constituent volumes.
2. Suspend I/O activity simultaneously across all volumes in the consistency group.
3. Create the snapshots.
4. Resume normal I/O activity across all volumes.

The XIV Storage System manages these suspend and resume activities for all volumes within the consistency group.

Consistency: Additional mechanisms or techniques, such as those provided by the Microsoft Volume Shadow Copy Service (VSS) framework, might still be required to maintain full application consistency from a host system perspective. For more information, see *IBM XIV Storage System Business Continuity Functions*, SG24-7759.

2.6.3 Storage pool relationships and rules

Storage pools facilitate the administration of relationships among logical volumes, snapshots, and consistency groups.

Storage pools have the following size considerations:

- For a fully populated XIV Storage System equipped with 1 TB, 2 TB, 3 TB, or 4 TB drives, the storage pool size can range from 17 GB (the minimum size that can be assigned to a logical volume) to 80 TB (with 1 TB disk drives), 161 TB (with 2 TB disk drives), 243 TB (with 3 TB disk drives), or 325 TB (with 4 TB disk drives),
- For partially populated XIV Storage System configurations, the maximum hard storage pool size is always equal to the maximum usable capacity of the configuration or 80 TB (with 1 TB disk drives), 161 TB (with 2 TB disk drives), 243 TB (with 3 TB disk drives), or 325 TB (with 4 TB disk drives), whichever is less.
- *Snapshot reserve capacity* is defined within each storage pool and is effectively *maintained separately* from logical, or master, volume capacity. The same principles apply for thinly provisioned storage pools, with the exception that space is not *guaranteed* to be available for snapshots because of the potential for hard space depletion.

The following principles govern the relationships between logical entities within the storage pool:

- An XIV Storage System LUN or logical volume can have multiple independent snapshots. This logical volume is also known as a *master volume*.
- A master volume and all of its associated snapshots are always in the same storage pool.
- A volume can only be part of a single consistency group and a single storage pool.
- All volumes of a consistency group must be in the same storage pool.
- The space allocated for a storage pool can be dynamically changed by the storage administrator:
  - The storage pool can be increased in size.
  - The storage pool can be decreased in size. It is limited only by the space that is used by the volumes and snapshots that are defined within that storage pool.
- The designation of a storage pool as a regular pool or a thinly provisioned pool can be dynamically changed even for existing storage pools. Thin provisioning is described in 2.7, “Capacity allocation and thin provisioning” on page 37.
- The storage administrator can relocate logical volumes between storage pools without any limitations, if there is sufficient free space in the target storage pool:
  - If necessary, the target storage pool capacity can be dynamically increased before volume relocation, assuming sufficient deallocated capacity is available in the system.
  - When a logical volume is relocated to a target storage pool, sufficient space must be available for all of its snapshots to be in the target storage pool as well.

Important:

- When moving a volume into a storage pool, the size of the storage pool is not automatically increased by the size of the volume. When removing a volume from a storage pool, the size of the storage pool does not decrease by the size of the volume.
- The system defines capacity using decimal metrics. 1 GB is 1 000 000 000 bytes using decimal metrics. By contrast, 1 GiB is 1 073 741 824 bytes using binary metrics.

The following principles apply to snapshots:

- Snapshots are structured in the same manner as logical, or master, volumes.

Snapshot reserve: The system preemptively deletes snapshots if the snapshots fully use the allocated space of the storage pool. Therefore, you must ensure that adequate space is allocated for snapshot reserve when defining a storage pool.

The snapshot reserve must be a minimum of 34 GB.

- Snapshots are automatically deleted only when there is inadequate physical capacity available within the context of each storage pool. This process is managed by a snapshot deletion priority scheme. Therefore, when the capacity of a storage pool is exhausted, only the snapshots that are in the affected storage pool are deleted in order of the deletion priority.

Important:
Snapshot deletion occurs automatically and older snapshots will be deleted without warning whenever a new snapshot is taken and space is insufficient to hold new volume or snapshot data.

To control this behavior, monitor the following events:

- STORAGE_POOL_VOLUME_USAGE_INCREASED
- STORAGE_POOL_SNAPSHOT_USAGE_INCREASED

Then, increase the pool size accordingly.

2.7 Capacity allocation and thin provisioning

Thin provisioning is a central theme of the virtualized design of the XIV Storage System because it uncouples the virtual, or apparent, allocation of a resource from the underlying hardware allocation.

The XIV Storage System implementation of thin provisioning provides these benefits:

- Capacity associated with specific applications or departments can be dynamically increased or decreased per the demand imposed at a specified point in time, without necessitating an accurate prediction of future needs. Physical capacity is only committed to the logical volume when the associated applications execute writes, as opposed to when the logical volume is initially allocated.

- Because the total system capacity is designed as a globally available pool, thinly provisioned resources share the “buffer” of free space. This results in highly efficient aggregate capacity use without pockets of inaccessible unused space.

  With the static, inflexible relationship between logical and physical resources commonly imposed by traditional storage subsystems, each application’s capacity must be managed and allocated independently. This situation often results in a large percentage of the total system capacity remaining unused because the capacity is confined within each volume at a highly granular level.
- Capacity acquisition and deployment can be more effectively deferred until actual application and business needs demand additional space, in effect facilitating an on-demand infrastructure.

Logical and actual volume sizes

The physical capacity that is assigned to traditional volumes is equivalent to the logical capacity presented to hosts. This situation does not have to be the case with the XIV Storage System thin provisioning. For a specified XIV Storage System logical volume, there are effectively two associated sizes, the logical volume size and the actual volume size. The physical capacity allocated for the volume is not static, but it increases as host writes fill the volume.

Logical volume size

The *logical volume size* is the size of the logical volume that is observed by the host, as defined upon volume creation or as a result of a resizing command. The storage administrator specifies the volume size in the same manner regardless of whether the volume’s storage pool is a thin pool or a regular pool.

The volume size is specified in one of two ways, depending on units:

- In terms of gigabytes: The system allocates the soft volume size as the minimum number of discrete 17 GB increments needed to meet the requested volume size.
- In terms of blocks: The capacity is indicated as a discrete number of 512-byte blocks. The system allocates the soft volume size used within the storage pool as the minimum number of discrete 17 GB increments needed to meet the requested size (specified in 512-byte blocks. However, the size that is reported to hosts is equivalent to the precise number of blocks defined.

The snapshot reserve capacity associated with each storage pool is a soft capacity limit. It is specified by the storage administrator, although it effectively limits the hard capacity used collectively by snapshots also.

Actual volume size

Tip: Defining logical volumes in terms of blocks is useful when you must precisely match the size of an existing logical volume on another system.

This value reflects the total size of volume areas that were written by hosts. The actual volume size is not controlled directly by the user and depends only on the application behavior. It starts from zero at volume creation or formatting and can reach the logical volume size when the entire volume has been written. Resizing of the volume affects the logical volume size, but does not affect the actual volume size.

The actual volume size reflects the physical space used in the volume as a result of host writes. It is discretely and dynamically provisioned by the system, not the storage

administrator. The discrete additions to actual volume size can be measured in either of these two ways:

- By considering the allocated space, which reflects the physical space used by the volume in 17 GB increments
- By considering the used space, which reflects the physical space used by the volume in 1 MB partitions

In both cases, the upper limit of this provisioning is determined by the logical size assigned to the volume:

- Capacity is allocated to volumes by the system in increments of 17 GB because of the underlying logical and physical architecture. There is no smaller degree of granularity than 17 GB. For more information, see 2.4, “Logical system concepts” on page 26.
- Application write access patterns determine the rate at which the allocated hard volume capacity is used and, therefore, the rate at which the system allocates additional increments of 17 GB, up to the limit defined by the logical volume size. As a result, the storage administrator has no direct control over the actual capacity allocated to the volume by the system at any specified point in time.
- During volume creation, or when a volume has been formatted, there is *zero physical capacity assigned to the volume*. As application writes accumulate to new areas of the volume, the physical capacity allocated to the volume grows in increments of 17 GB and can ultimately reach the full logical volume size.
- Increasing the logical volume size *does not affect* the actual volume size.

Thinly provisioned storage pools

Where volumes are effectively thinly provisioned automatically by the system, storage pools can be defined by the storage administrator (when using the XIV Storage Management GUI) as either *regular* or *thinly provisioned.* When using the IBM XIV Storage System Command-Line Interface (XCLI), there is no specific parameter to indicate thin provisioning for a storage pool. You indirectly and implicitly create a storage pool as thinly provisioned by specifying a pool soft size greater than its hard size.

With a regular pool, the “host-apparent” capacity is *guaranteed* to be equal to the physical capacity reserved for the pool. The total physical capacity *allocated to* the constituent individual volumes and collective snapshots at a particular time within a regular pool reflects the current use by hosts because the capacity is dynamically used as required. However, the remaining unallocated space within the pool remains reserved for the pool and cannot be used by other storage pools.

In contrast, a thinly provisioned storage pool is not fully backed by hard capacity, meaning that the entirety of the logical space within the pool cannot be physically provisioned unless the pool is transformed first into a regular pool. However, benefits can be realized when physical space consumption is less than the logical space assigned because the amount of logical capacity assigned to the pool that is not covered by physical capacity is available for use by other storage pools.

Figure 2-7 on page 40 shows a regular storage pool and a thin pool with these characteristics:

- In the regular pool, the host system sees a 34 GB LUN, a 51 GB LUN, and a 68 GB LUN. The storage pool size is the total of all three LUNs, which is 153 GB. About 40% of this storage is used.
- In the thin pool, the host system sees the same three LUN sizes, and the total storage pool size is also 153 GB. The difference is that the total space corresponding to unused

portions of each LUN, 34 GB, 51 GB, and 68 GB in the XIV, is not dedicated to those three LUNs, but remains available for other storage purposes.

When a storage pool is created using thin provisioning, that pool is defined in terms of both a soft size and a hard size independently, as opposed to a regular storage pool in which these sizes are by definition equivalent.

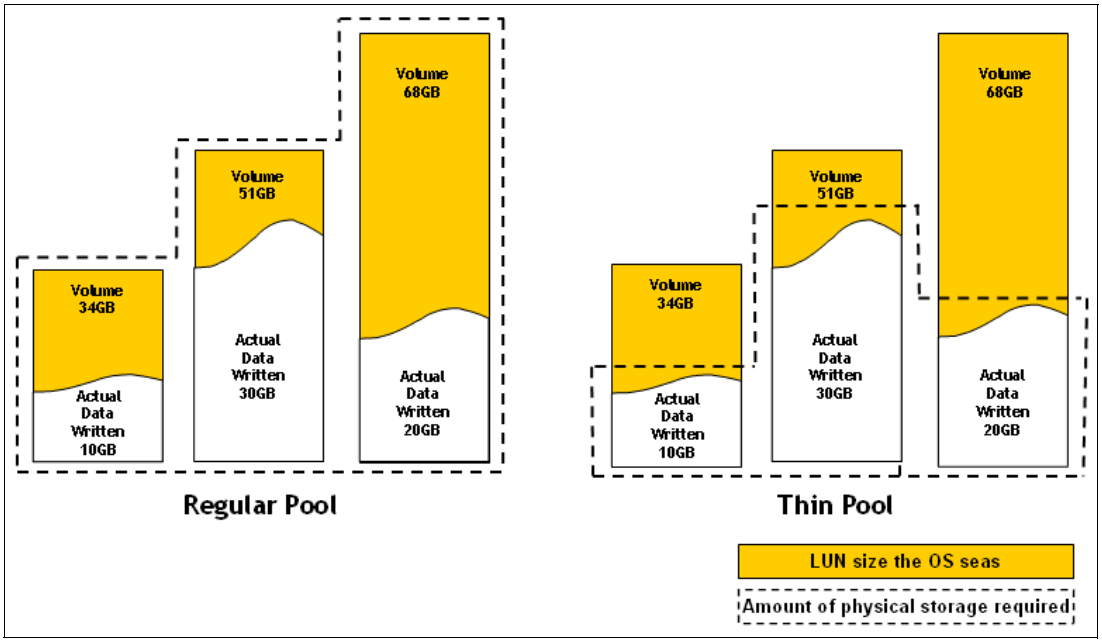


Figure 2-7 Regular pool versus thin pool

Hard pool size

Hard pool size is the maximum actual capacity that can be used by all the volumes and snapshots in the pool.

Thin provisioning of the storage pool maximizes capacity use in the context of a group of volumes, where the aggregate “host-apparent,” or soft, capacity assigned to all volumes surpasses the underlying physical, or hard, capacity allocated to them. This use requires that the aggregate space available to be allocated to hosts within a thinly provisioned storage pool must be defined independently of the physical, or hard, space allocated within the system for that pool.

Therefore, the storage pool hard size that is defined by the storage administrator limits the physical capacity that is available collectively to volumes and snapshots within a thinly provisioned storage pool. The aggregate space that is assignable to host operating systems is specified by the storage pool soft size.

Regular storage pools segregate the hard space reserved for volumes from the hard space used by snapshots by limiting the soft space allocated to volumes. However, thinly provisioned storage pools allow the total hard space to be used by volumes with no guarantee of preserving any hard space for snapshots.

Logical volumes take precedence over snapshots and can be allowed to overwrite snapshots, if necessary, as hard space is used. However, the hard space that is allocated to the storage pool that is unused (the incremental difference between the aggregate logical and actual *volume* sizes) can be used by snapshots in the same storage pool.

Careful management is critical to prevent hard space for both logical volumes and snapshots from being exhausted. Ideally, hard capacity use must be maintained under a certain threshold by increasing the pool hard size as needed.

Storage pools and snapshots:

- As described in "Storage pool relationships and rules" on page 36, storage pools control when and which snapshots are deleted when there is insufficient space assigned within the pool for snapshots.
- The soft snapshot reserve capacity and the hard space allocated to the storage pool are used only as changes occur to the master volumes or the snapshots themselves, not as snapshots are created.
- For a more detailed explanation of snapshot deletion priority, see *IBM XIV Storage System Business Continuity Functions*, SG24-7759.

Soft pool size

Soft pool size is the maximum logical capacity that can be assigned to all the volumes and snapshots in the pool.

Thin provisioning is managed for each storage pool independently of all other storage pools:

- Regardless of any unused capacity that might be in other storage pools, snapshots within a storage pool are deleted by the system according to the corresponding snapshot pre-set priority if the hard pool size contains insufficient space to create an additional volume or increase the size of an existing volume. (Snapshots are deleted only when a write occurs under those conditions, and not when allocating more space).
- As described in "Thinly provisioned storage pools" on page 39, the storage administrator defines both the soft size and the hard size of thinly provisioned storage pools and allocates resources to volumes within a storage pool without any limitations imposed by other storage pools.

The designation of a storage pool as a regular pool or a thinly provisioned pool can be dynamically changed by the storage administrator:

- When a regular pool needs to be converted to a thinly provisioned pool, the soft pool size parameter must be explicitly set in addition to the hard pool size, which remains unchanged unless updated.
- When a thinly provisioned pool must be converted to a regular pool, the soft pool size is automatically reduced to match the current hard pool size. If the combined allocation of soft capacity for existing volumes in the pool exceeds the pool hard size, the storage pool cannot be converted. This situation can be resolved if individual volumes are selectively resized or deleted or moved to another storage pool to reduce the soft space used.

System-level thin provisioning

The definitions of hard size and soft size apply at the subsystem level because it is necessary to allow the full system to be defined in terms of thin provisioning to achieve the full potential benefit previously described: The ability to defer deployment of additional capacity on an as-needed basis.

The XIV Storage System architecture allows you to define a global system capacity in terms of both a *hard system size* and a *soft system size*. When thin provisioning is not activated at the system level, these two sizes are equal to the system's physical capacity.

Hard system size

The *hard system size* represents the physical disk capacity that is available within the XIV Storage System. Obviously, the system's hard capacity is the upper limit of the aggregate hard capacity of all the volumes and snapshots and can be increased only by installing new hardware components in the form of individual modules (and associated disks) or groups of modules.

There are conditions that can *temporarily* reduce the system's hard limit. For more information, see 2.10.2, "Preserving data redundancy: Rebuilding and redistributing" on page 51.

Soft system size

The *soft system size* is the total, "global," and logical space available for all storage pools in the system. When the soft system size exceeds the hard system size, it is possible to logically provision more space than is physically available, allowing the benefits of thin provisioning of storage pools and volumes to be realized at the system level.

The soft system size limits the soft size of all volumes in the system and has the following attributes:

- It is not related to any direct system attribute and can be defined to be larger than the hard system size if thin provisioning is implemented. The storage administrator cannot set the soft system size.

  Storage pools: If the storage pools within the system are thinly provisioned, but the soft system size does not exceed the hard system size, the total system hard capacity cannot be filled until all storage pools are regularly provisioned. Therefore, it is best to define all storage pools in a non-thinly provisioned system as regular storage pools.

- The soft system size is a purely logical limit. However, you must exercise care when the soft system size is set to a value greater than the maximum potential hard system size. It must be possible to upgrade the system's hard size to be equal to the soft size. Therefore, defining an unreasonably high system soft size can result in full capacity depletion. It is for this reason that defining the soft system size is not within the scope of the storage administrator role.
- If it is necessary to increase the soft system size beyond the maximum hard system size for a particular XIV Storage System model, there is a procedure that IBM performs that can accomplish this task. Contact your IBM Technical Advisor or IBM sales team for more details.

There are conditions that might *temporarily* reduce the system's soft limit. For more information, see 2.10.2, "Preserving data redundancy: Rebuilding and redistributing" on page 51.

Depletion of hard capacity

Using thin provisioning creates the inherent danger of exhausting the available physical capacity. If the soft system size exceeds the hard system size, the potential exists for applications to fully deplete the available physical capacity.

Important: Upgrading the system beyond the full 15 modules in a single frame is currently not supported.

Snapshot deletion

As mentioned previously, snapshots in regular storage pools can be automatically deleted by the system to provide space for newer snapshots. For thinly provisioned pools, snapshots can be deleted to free more physical space for volumes.

Volume locking

If more hard capacity is still required after all the snapshots in a thinly provisioned storage pool have been deleted, *all the volumes in the storage pool are locked, preventing any additional consumption of hard capacity.* There are two possible behaviors for a locked volume: read only (the default behavior) or no I/O at all. In either case, your applications stop in a way that is not predictable.

Important: Volume locking prevents writes to all volumes in the storage pool.

The thin provisioning implementation in the XIV Storage System manages space allocation within each storage pool so that hard capacity depletion in one storage pool never affects the hard capacity available to another storage pool. There are both advantages and disadvantages:

- Because storage pools are independent, thin provisioning volume locking on one storage pool never cascades into another storage pool.
- Hard capacity cannot be reused across storage pools, even if a certain storage pool has free hard capacity available. This limitation can lead to a situation where volumes are locked because of the depletion of hard capacity in one storage pool while there is available capacity in another storage pool. It is still possible for the storage administrator to intervene to redistribute hard capacity.

2.8 Flash caching architecture

Starting with the XIV System software Version 11.1.0, the XIV Storage System Gen3 supports an optional flash caching feature that significantly improves performance for small block, random-read I/O workloads.

2.8.1 Flash caching overview

XIV flash caching is implemented as an extension of the primary cache layer. Flash cache is used exclusively as a read cache. Each module supports 400 GB cache, which allows the caching layer to be expanded up to 6 TB. Starting with software Version 11.4, a flash cache of 800 GB per module for a total of 12 TB in a fully populated system with 4 TB drives can be used optionally to further boost performance.

With flash caching, as illustrated in Figure 2-8 on page 44, there is no need to relocate data. Because flash cache is used as read cache only, when data in cache is no longer accessible, it can simply be dropped and replaced by more relevant data.

This approach allows for the flash cache to be used effectively in highly dynamic environments where data patterns are constantly changing. To have a holistic solution, the caching architecture must also deliver good write performance.

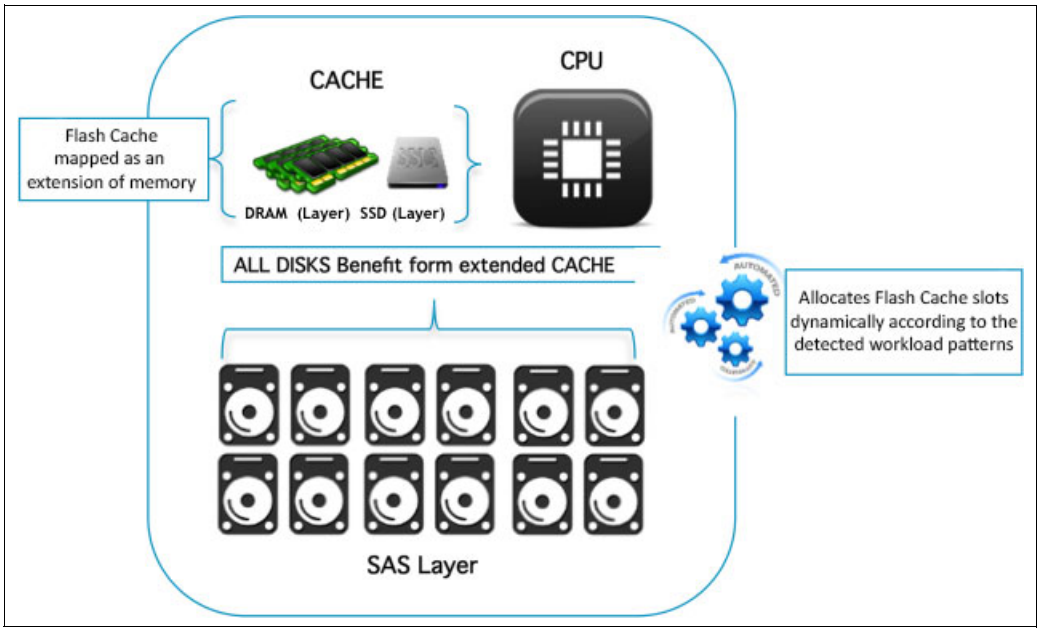

CACHE
CPU
Flash Cache mapped as an extension of memory
DRAM (Layer) SSD (Layer)
ALL DISKS Benefit form extended CACHE
Allocates Flash Cache slots dynamically according to the detected workload patterns
SAS Layer

Figure 2-8 XIV flash caching overview

The caching algorithm is embedded in the XIV System software (firmware) and makes the flash cache integration completely transparent to the user or storage administrator. No tuning of the cache is required to get the potential performance boost.

The flash caching feature is supported automatically by the XIV unique caching algorithm, which dynamically adapts to detected I/O request patterns.

2.8.2 Flash caching algorithm and concepts

Caching operations are split into two groups, main and extended:

- The *main cache* handles host write I/Os and then destages them directly to the disk drive.
- The *extended cache* handles the caching of random read miss operations less than 64 KB. Sequential read prefetches (larger than 64 KB) are handled in main dynamic random access memory (DRAM) cache.

Flash caching: For more details about flash caching, see the Redpaper publication, *Solid-State Drive Caching Implementation in the IBM XIV Storage System*, REDP-4842.

A flash cache map is built as read misses occur in the DRAM cache. The process, known as *flash cache learning,* is depicted in Figure 2-9 on page 45.

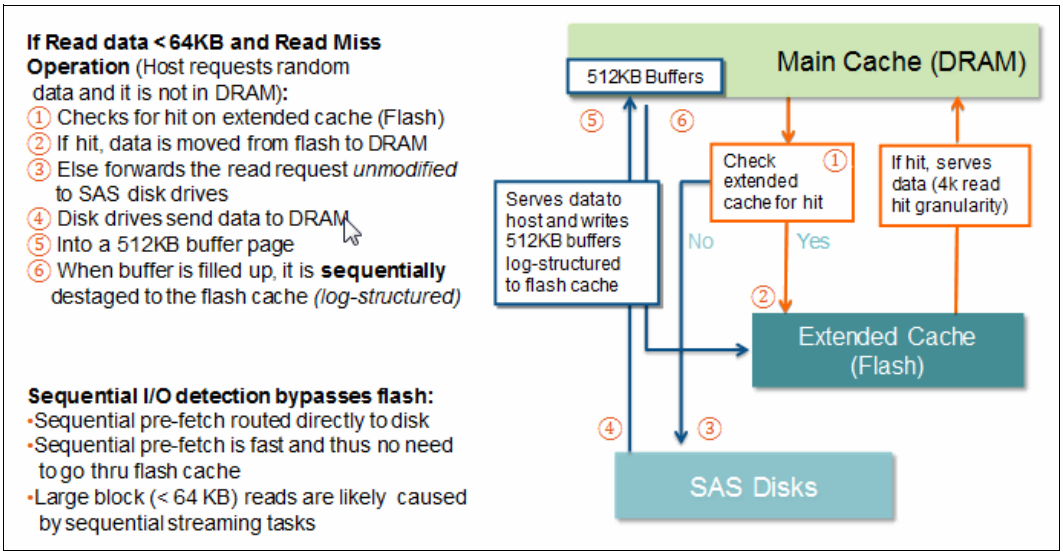

If Read data < 64KB and Read Miss Operation (Host requests random data and it is not in DRAM):
① Checks for hit on extended cache (Flash)
② If hit, data is moved from flash to DRAM
③ Else forwards the read request unmodified to SAS disk drives
④ Disk drives send data to DRAM
⑤ Into a 512KB buffer page
⑥ When buffer is filled up, it is sequentially destaged to the flash cache (log-structured)
Sequential I/O detection bypasses flash:
•Sequential pre-fetch routed directly to disk
•Sequential pre-fetch is fast and thus no need to go thru flash cache
•Large block (< 64 KB) reads are likely caused by sequential streaming tasks
Main Cache (DRAM)
512KB Buffers
Check extended cache for hit
If hit, serves data (4k read hit granularity)
Serves data to host and writes 512KB buffers log-structured to flash cache
No
Yes
Extended Cache (Flash)
SAS Disks

Figure 2-9 XIV flash cache learning

The cache node immediately checks the extended cache for the requested I/O. If the requested I/O exists in the extended flash cache, it is served to the host through the main cache. The I/O operation is now complete and is recorded as a flash cache read hit.

If the operation results in a true read miss (not in the DRAM cache and not in extended flash cache), the request is forwarded in an unmodified state to the disk drive (SAS layer). The I/O is retrieved from the disk drive and served to the host through the main cache. From a host perspective, the I/O operation is now complete and is recorded as a read miss. The related pages are copied into reserved buffers in the main cache.

Important: Any read larger than 64 KB bypasses the extended flash cache.

When the buffer reaches 512 KB, it is written sequentially to the flash cache as a log-structured write. This method helps to prolong the life of the flash cache.

Note: The XIV is able to retain the data in the flash cache between system restarts and code upgrades.

XIV Storage System Software Version 11.2 introduced improved flash caching algorithms, which provide a performance boost of up to 4.5 times over systems without flash cache for random database-type workloads. This boost is accomplished by storing and computing all flash cache-related data integrity checking tasks in DRAM rather than on the flash cache.

2.9 Encryption for data at rest

IBM XIV Storage System Gen3 Model 214 with software Version 11.4.0, for machine types 2810 and 2812, introduced industry-standard encryption for data at rest while avoiding a performance impact. All capacity points, 1 TB, 2 TB, 3 TB, 4TB and 6 TB, are supported using self-encrypting hard drives (SED) and software-based, encryption-ready flash cache. Implemented with AES 256-bit keys, encryption can be done with data already on disk (hot encryption). It is nondisruptive and takes only minutes.

Key management is accomplished with Tivoli Key Lifecycle Manager or its replacement, IBM Security Key Lifecycle Manager, which offers production-ready key management and is based on the Key Management Interoperability Protocol (KMIP) Version 1.0. This method provides centralized and simplified key management and the separation of key storage from data storage.

Figure 2-10 shows the initial configuration of encryption in XIV.

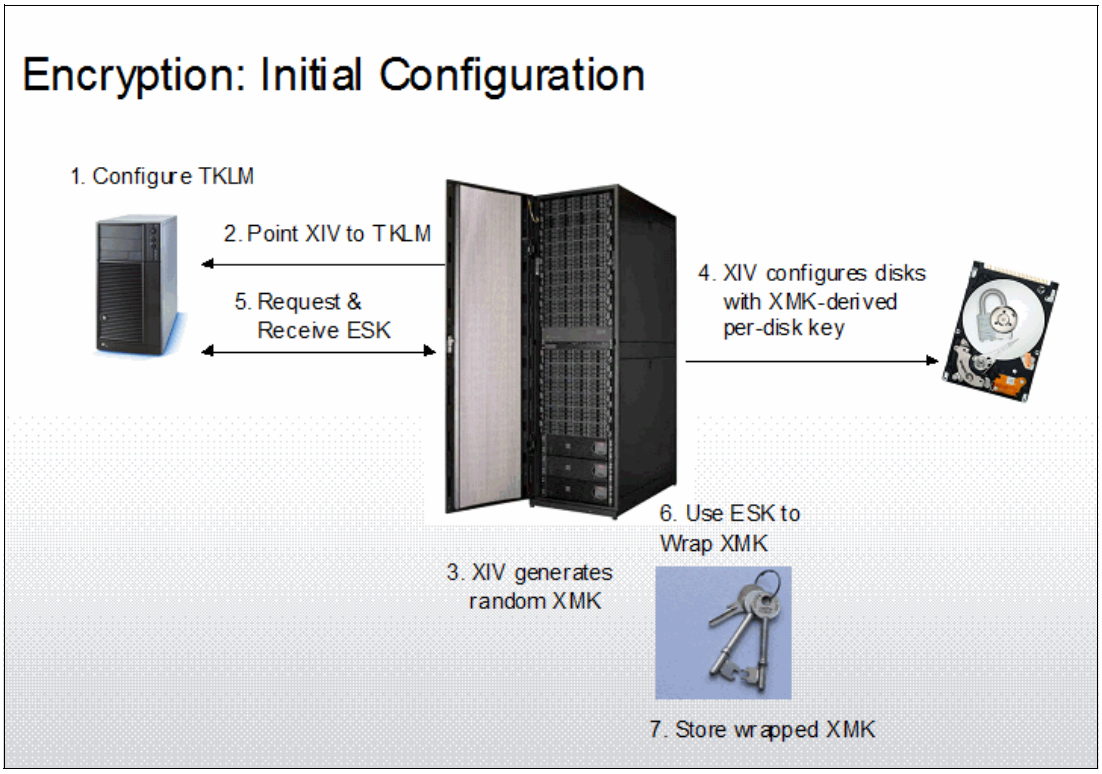

Encryption: Initial Configuration
1. Configure TKLM
2. Point XIV to TKLM
5. Request & Receive ESK
4. XIV configures disks with XMK-derived per-disk key
6. Use ESK to Wrap XMK
3. XIV generates random XMK
7. Store wrapped XMK

Figure 2-10 Encryption: Initial configuration

For more information, see the IBM Redpaper, *XIV Security with Data-at-Rest Encryption*, REDP-5047.

2.10 Reliability, availability, and serviceability (RAS)

The unique modular design and logical topology of the XIV Storage System fundamentally differentiates it from traditional, monolithic systems. This architectural divergence extends to the exceptional reliability, availability, and serviceability aspects of the system. The XIV also incorporates autonomic and proactive monitoring and self-healing features. These features can take preventive measures to preserve data redundancy before a component malfunction occurs and automatically restore the system to full redundancy within minutes of a hardware failure.

For more information about the XIV Storage System parallel modular architecture, see 2.2, “Parallelism” on page 21.

2.10.1 Resilient architecture

As with any enterprise class system, redundancy pervades every aspect of the XIV Storage System, including the hardware, internal operating environment, and the data itself. However, the design elements, including the distribution of volumes across the whole of the system, in combination with the loosely coupled relationship between the underlying hardware and software elements, empower the XIV Storage System to realize unprecedented resiliency. The resiliency of the architecture encompasses not only high availability, but also excellent maintainability, serviceability, and performance under ideal conditions resulting from planned or unplanned changes to the internal hardware infrastructure, such as the loss of a module.

Availability

The XIV Storage System maximizes continuous operation and minimizes the performance degradation associated with nondisruptive planned and unplanned events, while providing the capability to preserve the data in a disaster.

High reliability

The XIV Storage System not only withstands individual component failures by quickly and efficiently reinstating full data redundancy, it also monitors permanently and phases out individual components before data redundancy is compromised. We describe this in “Proactive phase out and self-healing mechanisms” on page 58. The collective high-reliability provisions incorporated within the system constitute multiple layers of protection from unplanned outages and minimize the possibility of related service actions.

Maintenance freedom

Although the potential for unplanned outages and associated corrective service actions are mitigated by the reliability attributes inherent in the system design, the XIV Storage System autonomic features minimize the need for storage administrators to conduct non-preventative maintenance activities that are purely reactive in nature. This is done by adapting to potential issues before they are manifested as a component failure. The continually restored redundancy, along with the self-healing attributes of the system, effectively enable maintenance activities to be decoupled from the instigating event (such as a component failure or malfunction) and safely carried out according to a predefined schedule.

In addition to the system’s diagnostic monitoring and autonomic maintenance, a proactive and systematic approach to maintenance is augmented because the entirety of the logical topology is continually preserved, optimized, and balanced according to the physical state of the system.

The modular system design also expedites the installation of any replacement or upgraded components, while the automatic and transparent data redistribution across all resources eliminates the downtime, even in the context of individual volumes, associated with these critical activities.

High availability

The rapid restoration of redundant data across all available drives and modules in the system during hardware failures, and the equilibrium resulting from the automatic redistribution of data across all newly installed hardware, are fundamental characteristics of the XIV Storage System architecture. These capabilities minimize exposure to cascading failures and the associated loss of access to data.

Consistent performance

The XIV Storage System can adapt to the loss of an individual drive or module efficiently and with relatively minor impact compared to monolithic architectures.

Although predefined traditional monolithic systems employ an N+1 hardware redundancy scheme, the XIV Storage System uses a grid topology. This topology can sustain a component failure and maximize consistency and transparency from the perspective of attached hosts. The potential impact of a component failure is vastly reduced because each module in the system is responsible for a relatively small percentage of the system's operation.

A controller failure in a typical N+1 system results in a dramatic (up to 50%) reduction of available cache, processing power, and internal bandwidth. Conversely, the loss of a module in the XIV Storage System translates to only 1/15th of the system resources and does not compromise performance nearly as much as the same failure with a typical architecture.

For more information, see Chapter 6, "Performance" on page 287.

The XIV Storage System also incorporates innovative provisions to mitigate isolated disk-level performance anomalies through *redundancy-supported reaction*. This is described in "Redundancy-supported reaction" on page 59. Also see "Flexible handling of dirty data" on page 59.

Disaster recovery

Enterprise class environments must account for the possibility of the loss of both the system and all of the data as a result of a disaster. The XIV Storage System includes the provision for remote mirror function as a fundamental component of the disaster recovery strategy. For more information about the remote mirror function in XIV, see *IBM XIV Storage System Business Continuity Functions*, SG24-7759.

Write path redundancy

Data arriving from the hosts is temporarily placed in two separate caches before the XIV Storage System acknowledges that the write is received and begins to write the data to the disk drives in separate modules. This design guarantees that the data is always protected against possible failure of individual modules, even before the data is written to the disk drives.

Figure 2-11 illustrates the path that is taken by a write request as it travels through the system.

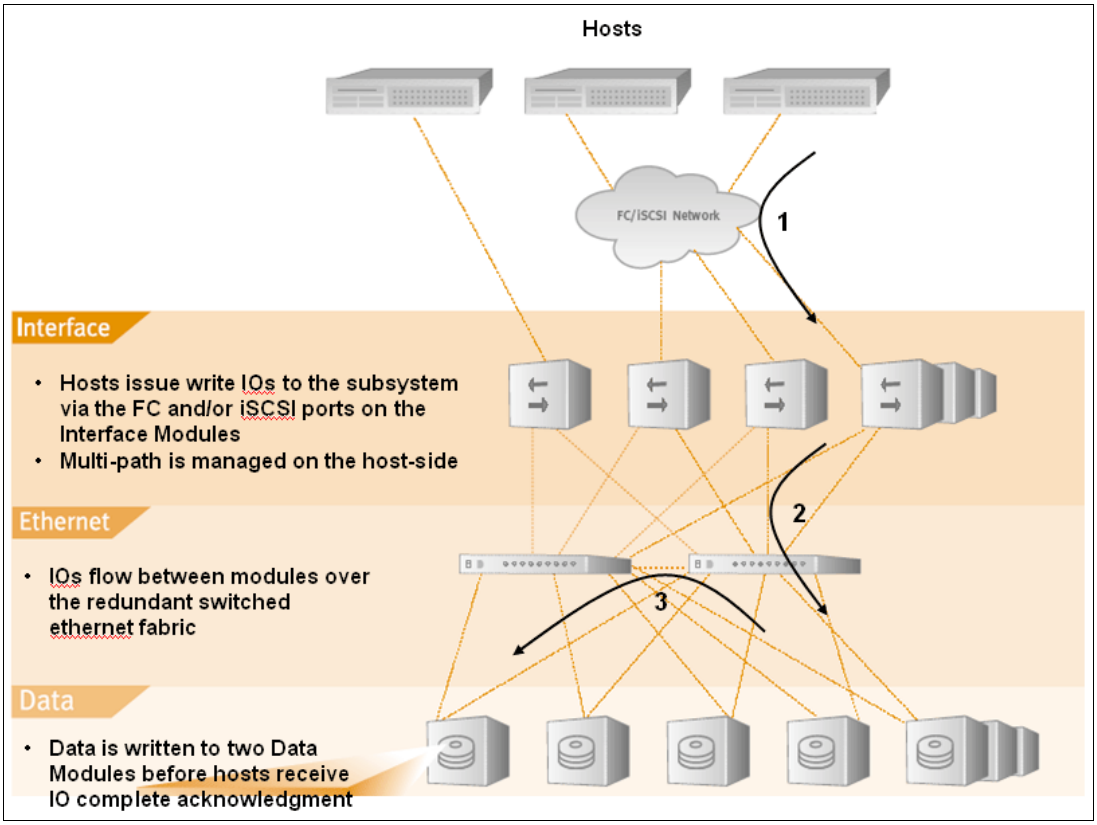

Hosts
FC/iSCSI Network
1
Interface
• Hosts issue write IOs to the subsystem via the FC and/or iSCSI ports on the Interface Modules
• Multi-path is managed on the host-side
2
Ethernet
• IOs flow between modules over the redundant switched ethernet fabric
3
Data
• Data is written to two Data Modules before hosts receive IO complete acknowledgment

Figure 2-11 Write path

Note: The diagram is intended to be viewed as a conceptual topology. Therefore, do not interpret the specific numbers of connections and so on as literal depictions. Also, for our purposes, the interface modules are depicted on a separate level from the data modules. However, in reality, the interface modules also function as data modules.

The following numbers correspond to the numbers in Figure 2-11:

1. A host sends a write request to the system. Any of the interface modules that are connected to the host can service the request because the modules work in an active-active capacity. The XIV does not balance the load for write requests across the different interface modules. Load balancing must be implemented by host administrators by using multipathing and by the SAN administrators, using correct zoning, to equally distribute the host requests among all interface modules.

   For more information about host attachment, see the Redbooks publication, *XIV Storage System: Host Attachment and Interoperability,* SG24-7904.

2. The interface module uses the system configuration information to determine the location of the primary module that houses the referenced data. This can be either an interface module, including the interface module that received the write request, or a data module. The data is written only to the local cache of the primary module.

3. The primary module uses the system configuration information to determine the location of the secondary module that houses the copy of the referenced data. Again, this module can be either an interface module or a data module, but it is not the same as the primary module. The data is redundantly written to the local cache of the secondary module.

After the data is written to cache in both the primary and secondary modules, the host receives an acknowledgment that the I/O is complete, which occurs independently of the data being destaged to physical disk.

System shutdown and power-on sequence

When the XIV Storage System detects that all power is lost, it waits for 30 seconds to ensure that this outage is not a temporary power outage. During this time, it runs on three uninterruptible power supply (UPS) units that are part of every XIV Storage System configuration. The XIV Storage System needs only two of these units to be operational. The third UPS unit is redundant. If power is not restored within 30 seconds, the XIV Storage System executes a graceful shutdown sequence. There is enough battery power in the XIV Storage System to execute such a shutdown sequence twice.

Because of the XIV Storage System grid topology, a system shutdown event essentially entails the graceful shutdown of all modules within the system. Each module can be thought of as an independent entity that is responsible for managing the destaging of all data in cache that has not already been written to the disk drive. The data in cache that must be written to disk within each module consists of equal parts primary and secondary copies of data, *but never contains both primary and secondary copies of the same data.*

Write cache protection

Each module in the XIV Storage System contains a local, independent space reserved for caching operations within its system memory.

Destage: The system memory space is reserved for write operations. However, the proximity of the cache and the disk drives, with the enforcement of an upper limit for data that has not been destaged and is enforced on a per-drive basis ensures that the full destage process occurs while operating under battery power.

Important: The extended flash cache layer is non-volatile memory, and all data in flash cache is protected for either scheduled or non-scheduled system shutdown. The XIV Storage System does not use flash cache for write. All write I/Os are staged and mirrored out of main cache (DRAM layer) only. The writes are flushed from main cache to the disk drive layer as part of the normal write destaging. However, on shutdown, the primary cache (DRAM) related metadata is dumped to the flash cache. Upon the next XIV startup, the metadata is read back and validated for correctness.

For more information, see the Redpaper publication, *Solid-State Drive Caching Implementation in the IBM XIV Storage System*, REDP-4842.

Shutdown trigger events

The system executes the graceful shutdown sequence under either of these conditions:

- The battery charge remaining in two or more UPS units hits a certain threshold, which is conservatively set just above the amount of battery power required to fully destage data from cache to disk and gracefully shutdown the entire system *twice.*
- The system detects the loss of external power for more than 30 seconds.

Power-on sequence

Upon startup, the system verifies that the battery charge levels in all UPS units exceed the threshold necessary to guarantee that a graceful shutdown can occur twice. If the charge level is inadequate, the system halts the startup process until the charge level reaches the minimum required threshold.

Important: If there is a complete power loss in the data center, the XIV *automatically* powers up when power is reapplied. If this is not the behavior you want, contact IBM technical support to learn how to disable this feature.

2.10.2 Preserving data redundancy: Rebuilding and redistributing

As described in “Data distribution algorithms” on page 23, the XIV Storage System dynamically maintains the pseudo-random distribution of data across all modules and disks while ensuring that two copies of data exist at all times when the system reports *Full Redundancy*. When there is a change to the hardware configuration as a result of adding new hardware or a failed component, data must be rebuilt if necessary and redistributed to the new pseudo-random distribution. Either way, the XIV Storage System data redundancy mechanisms must accommodate the change.

These redundancy mechanisms are sometimes referred to as *autonomic* because these procedures are self-managing. They happen without any storage administrator intervention.

When a disk drive or a module is replaced during normal maintenance, or additional data storage capacity is added to an existing XIV System configuration, the XIV System moves existing primary and secondary data partitions to the new hardware. This process occurs in such a way that results in an even distribution of partitions across all disk drives and modules. At no time during this process does the XIV System become non-redundant. This distribution process consists of the following activities:

- Creation of a target data distribution
- Initiation of the redistribution of the redundant data according to the new target data distribution, known as *redistribution*

Redistribution is further described in “Redistribution after adding modules or drives” on page 57.

When the full redundancy of data is compromised because of a disk drive or module failure, the XIV Storage System immediately identifies the non-redundant partitions and begins the rebuild process. The rebuild process is similar to the redistribution process, and consists of the following activities:

- Creation of a target data distribution
- Copying of the non-redundant partitions and writing them according to the new target distribution, known as *rebuilding*
- Simultaneously initiation of the redistribution of the redundant data according to the new target data distribution, known as *redistribution*

Note: After an XIV Storage System component failure, rebuild and redistribution begin immediately and at the same time.

The sections that follow describe XIV data rebuild and redistribution.

Rebuilding after a failure

When a disk drive or a module fails, the XIV data distribution algorithms automatically identify the non-redundant partitions and then begins the *rebuilding* process. The XIV Storage System then makes copies of just those non-redundant partitions and places this data in reserved areas of the disk drives that exist for this circumstance. During this rebuilding period, the XIV Storage System is not redundant. Therefore, this is a self-healing phase performed to restore the system to full redundancy. This phase is performed as a high priority.

Note: XIV Storage System Software Version 11.2 introduced enhancements that significantly reduce rebuild times by as much as 50% over previous versions. An enhanced algorithm, designated as disk Quality of Service (QoS) introduced in v 11.5.1 has further improved rebuild times by introducing a mechanism that allows doing more rebuild (and scrubbing) IOPS at times when the system is not under stress (see Figure 2-13 on page 54).

Table 2-2 shows both the disk and module rebuild times for a fully allocated 15 module XIV Storage System (Version 11.3) under significant workload.

Table 2-2 Rebuild times for a 15 module Gen3 XIV Storage System running software Version 11.3

Rebuild type	2 TB drives	3 TB drives	4 TB drives
Disk	26 minutes	39 minutes	52 minutes
Module	269 minutes	420 minutes	570 minutes

Important: Table 2-2 illustrates the rebuild times that are possible with XIV under load. Actual rebuild times depend on several factors:

- Numbers of modules
- Type of failure (disk or module)
- Host-generated workload
- Amount of data written to the system

The rebuild times that you experience in your environment depend on these factors.

To illustrate this rebuild scenario, we purposely failed module 13, as shown in Figure 2-12 on page 53. Notice the status bar in the lower-right corner of the GUI that shows the completed percentage of the rebuilding progress. The percentage completed is in light green and is reported in our example as 30% complete. The percentage that is yet to be completed is in yellow. Remember that the redistribution process is also running, even though the XIV Storage Management GUI reports only the rebuilding.

Figure 2-12 Rebuilding status in the XIV Storage System GUI

If there is a disk failure in a fully configured XIV Storage System, there are 168 disks reading because there is no data that is not redudant on the other disks within the same module as the failed disk. Concurrently, there are 179 disks writing to preserve full data distribution.

If there is a module failure, the copies of non-redundant data are read from all of the remaining modules in the system because none of the disks within a particular module contain the secondary copies of data on any of the disks in the module. Therefore, during a rebuild that results from a module failure in a fully configured XIV Storage System, there are 168 disks (180 disks in the system minus 12 disks in a module) reading and 168 disks writing, concurrently.

Important: The rebuilding phase is a high-priority, self-healing process to restore XIV to full redundancy. At the end of the rebuilding period, the XIV Storage System is again fully redundant.

The XIV Storage System rebuild process has the following characteristics:

- The rebuild of data is many times faster than conventional RAID array rebuilds and can complete in a short period for a fully provisioned system:
  - Statistically, the chance of exposure to data loss or a cascading hardware failure is minimized because of both the short rebuild time required and the low I/O workload imposed on any particular disk.

Rebuilding conventional RAID arrays can take many hours to complete, depending on the type of the array, the number of drives, and the ongoing host-generated transactions to the array. The stress imposed on the drives of conventional RAID arrays during rebuilds is high, raising the risk of additional hardware failure of the RAID member disks.

In XIV, a typical rebuild is measured in minutes, not hours. For example, a fully populated 15-module XIV system with the usable capacity of 240 TB typically rebuilds in less than 40 minutes. For a more information about rebuild times under different scenarios, see Table 2-2 on page 52.

– The rebuild process can complete 25% to 50% more quickly for systems that are not fully provisioned, which equates to a rebuild completion in as little as 15 minutes. The rebuild time is linear and decreases with the amount of written data.

- The disk rebuild QoS concept implemented in v 11.5.1 has further improved rebuild times by introducing a mechanism that allows doing more rebuild (and scrubbing) IOPS at times when the system is not under stress. Refer to Figure 2-13

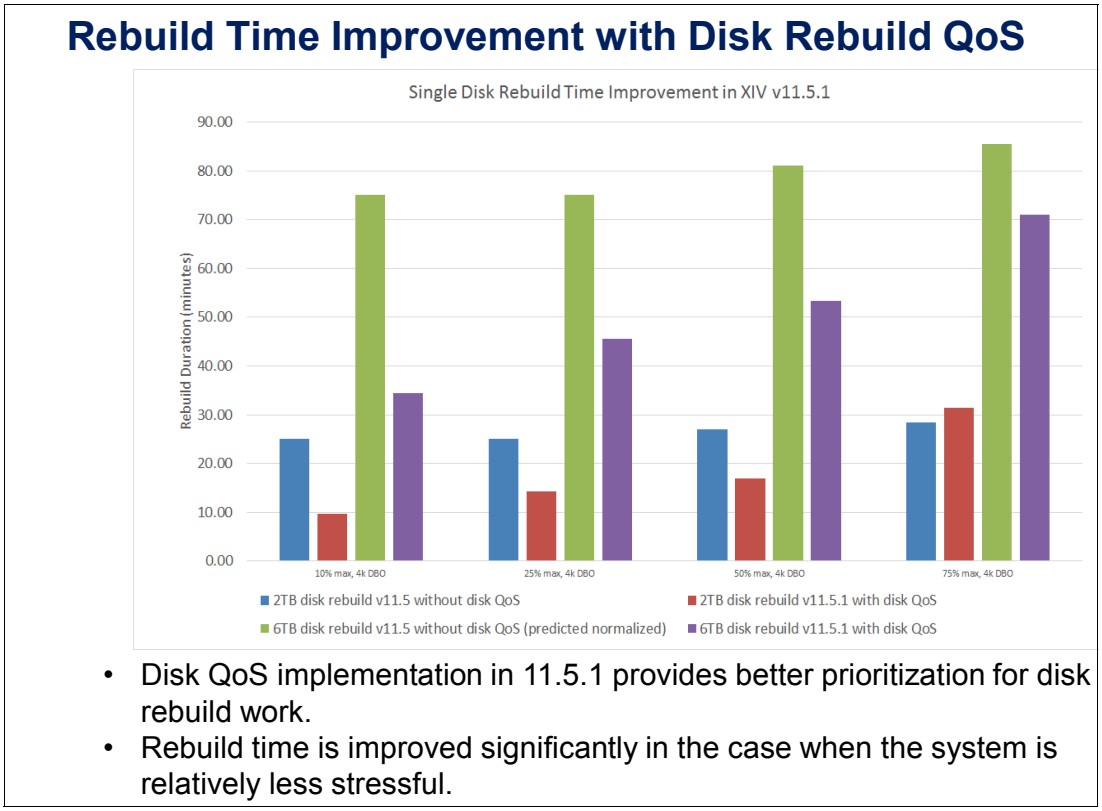

Rebuild Time Improvement with Disk Rebuild QoS
Single Disk Rebuild Time Improvement in XIV v11.5.1
Rebuild Duration (minutes)
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
10% max, 4k DBO
25% max, 4k DBO
50% max, 4k DBO
75% max, 4k DBO
2TB disk rebuild v11.5 without disk QoS
2TB disk rebuild v11.5.1 with disk QoS
6TB disk rebuild v11.5 without disk QoS (predicted normalized)
6TB disk rebuild v11.5.1 with disk QoS
• Disk QoS implementation in 11.5.1 provides better prioritization for disk rebuild work.
• Rebuild time is improved significantly in the case when the system is relatively less stressful.

Figure 2-13 Rebuild time improvement in XIV v 11.5.1.

- The XIV Storage System rebuilds only actual data. The number of drives participating in the rebuild is about 20 times greater than in most average-sized conventional RAID arrays. By comparison, the array rebuild workload is greatly dissipated, greatly reducing the relative impact on host performance:

  – In a conventional RAID array, the whole disk is re-created which often includes unused space.

  – Conventional RAID array rebuilds place many times the normal transactional load on the disks and substantially reduce effective host performance.

- The XIV Storage System maintains universally accessible reserve space on all disks in the system, as described in “Global spare capacity” on page 34.

  Standard dedicated spare disks used during a conventional RAID array rebuild might not be globally accessible to all arrays in the system.
- Hotspots are statistically eliminated because the XIV Storage System maintains access density equilibrium, which reduces the chances of isolated workload-induced failures.
- The XIV Storage System data distribution eliminates localized drive stress and associated additional heat generation.
- Modules intelligently send information to each other directly. There is no need for a centralized supervising controller to read information from one disk module and write to another disk module.
- All disks are monitored for errors, poor performance, or other signs that might indicate that a full or partial failure is impending.

  Dedicated spare disks in conventional RAID arrays are inactive, and therefore, unproven and potentially unmonitored, increasing the possibility for a second failure during an array rebuild.

Important: If an XIV Gen3 storage system is equipped with flash cache, the flash cache does not participate in the rebuild process because flash cache is only used as extended *read* cache. Therefore, the most recent written data must be taken from the disk drives.

For the same reason, a flash cache device failure does not initiate a rebuild or redistribution process.

For details, see 2.8, “Flash caching architecture” on page 43, or consult the Redpaper publication, *Solid-State Drive Caching Implementation in the IBM XIV Storage System*, REDP-4842.

Redistributing after a failure

The *redistribution* phase starts at the same time as the rebuilding phase after an XIV Storage System disk or module failure. However, the XIV Storage System GUI begins to report only on redistribution progress after the rebuilding phase is complete and the system has been restored to full redundancy.

In the redistributing phase, the XIV Storage System data partitions (both primary and secondary copies) continue to be rearranged into the new optimized pseudo-random target distribution. Redistribution is a low-priority self-tuning process. Notice the status bar in the lower-right corner of the XIV Storage Management GUI in Figure 2-14, which shows the progress of the redistributing process. The percentage complete is in bright green and the percentage that is yet to be complete is in light green.

Figure 2-14 Redistribution status in the XIV Storage System GUI

Important: Redistribution is a low-priority self-tuning process, during which the XIV Storage System is fully redundant.

Effects of a flash cache failure

If a flash cache device starts to fail or fails completely, it is phased out like any of the other XIV grid components. There is no data loss because the flash cache is not involved in any phase of the process of storing data or maintaining data redundancy. The flash caching algorithms populate the extended flash cache with copies of primary data partitions. If a flash cache device fails, no data is lost.

The flash cache degraded module continues to serve reads from its DRAM cache and large sequential reads from disk. All small read misses are redirected to the secondary partition copies, which are on modules with functioning flash cache. This behavior evenly distributes the extended flash caching duties of the module with the failed flash device to all other modules, therefore balancing the use of the remaining extended flash cache across all modules.

There is no rebuild or redistribution of data as a result of a flash device failure. All the primary partitions on the module with the failed flash device remain the primary partition. But because of the reduced extended flash cache, the specific module limits its caching duties.

Redistribution after adding modules or drives

The redistribution process that happens after a rebuild is the same process that results from the following events:

- The replacement of a failed disk drive or module following a rebuild, which is also known as a *phase in*
- The addition of one or more modules to upgrade the capacity of the system

Important: The client's storage administrators cannot perform a phase in to replace a failed disk or module. These operations must be performed by IBM technicians. Therefore, it is not a good idea to purposely pull an XIV Storage System disk drive to test the redundancy mechanisms, because storage administrators are not authorized to replace a drive back into the system.

The XIV Storage System data distribution has the following characteristics:

- The redistribution process is triggered by the phase in of a new drive or module and differs from a rebuild or phase out in the following ways:
  - The system does not need to create secondary copies of data to reinstate or preserve full data redundancy.
  - The concentration of data on each physical disk decreases.
- The redistribution is dependent on the writing capabilities to the new drive or module:
  - When a replacement module is phased in, there are 168 disks reading and 12 disks writing concurrently. Therefore, the time to completion is determined by the data throughput to the replacement module. Because the read frequency to the existing disks due to the redistribution process is low, a low impact on host performance during the process is guaranteed.
  - When a replacement disk is phased in, there are concurrently 179 disks reading and only one disk writing. In this case, the new drive determines the achievable throughput of the redistribution process. Again, the impact on host transactions is small and not noticeable.

Transient soft and hard system size

Capacity that is required for restoring non-redundant data during a rebuild is, depending on availability, obtained in this order:

1. *Deallocated system hard capacity*: Hard capacity that was not assigned to any storage pools at the time of the failure.
2. *Reserve spare capacity*: The system reserves enough capacity to sustain the consecutive, non-concurrent failure of three drives and an entire module. If it is already used, replacement hardware must be phased in to ensure that data redundancy can be restored.

When sufficient deallocated hard capacity is available, the system withholds allocating reserve spare space to complete the rebuild or phase-out process to provide additional protection. As a result, it is possible for the system to report a maximum soft size that is temporarily less than the allocated soft capacity. The soft and hard system sizes do not revert to the original values until a replacement disk or module is phased in, and the resulting redistribution process is completed.

Important: Although it is possible to resize or create volumes, snapshots, or storage pools while a rebuild is underway, do not perform these activities until the system completes the rebuild process and restores full data redundancy.

2.10.3 Exclusive additional functions for reliability and availability

This section describes other features that contribute to XIV Storage System reliability and availability.

Disaster recovery

All high availability SAN implementations must account for the contingency of data recovery and business continuance following a disaster, as defined by the organization's recovery point and recovery time objectives. The provision within the XIV Storage System to efficiently and flexibly create nearly unlimited snapshots, coupled with the ability to define consistency groups of logical volumes, constitutes integral elements of the data preservation strategy. In addition, the XIV data mirroring function facilitates excellent potential recovery point and recovery time objectives as a central element of the full disaster recovery plan. For more information, see *IBM XIV Storage System Business Continuity Functions*, SG24-7759.

Proactive phase out and self-healing mechanisms

A further enhancement to the level of reliability standards attained by the XIV Storage System entails self-diagnosis and early detection mechanisms that automatically phase out components before they actually fail.

In real systems, the failure rate is not constant over time, but rather increases with service life and duty cycle. By actively gathering component statistics to monitor this trend, the system ensures that components do not operate under conditions beyond an acceptable threshold of reliability and performance. Therefore, the XIV Storage System self-healing mechanisms increase the already high level of availability of the system even further and also safeguard critical operations, such as a rebuild from further component failures.

Total cost of ownership (TCO) is an important advantage of the XIV Storage System. The self-healing mechanisms make service actions less frequent. When service actions are necessary to replace drives or modules, the XIV has already automatically undergone rebuild and redistribution tasks to adapt to the changing configuration. After component replacement, the XIV Storage System automatically absorbs the new hardware into the new configuration. The XIV Storage System leads the industry in the amount of time that is required to adapt to the new configurations. All of these provisions together minimize maintenance time and costs and are key to the low TCO of the XIV.

Disk scrubbing

The XIV Storage System maintains a scrubbing algorithm that runs continuously as a background process scanning all disks for checksum errors. It assures data integrity by alerting corrective actions before user data becomes compromised.

Therefore, redundancy is not only implemented as part of the basic architecture of the system, but it is also continually monitored and restored as required. In summary, the data scrubbing process has the following attributes:

- Verifies the integrity and redundancy of stored data, even across mirrors
- Enables early detection of disk errors and their early recovery
- Runs as a background process on each module and all disks simultaneously
- Zeroes out partitions that are not allocated to the user data space

A complete scrubbing cycle takes about one week for a system with 1 TB disks, two weeks for a system with 2 TB disks, and so on.

Enhanced monitoring and disk diagnostic tests

The XIV Storage System continuously monitors the performance level and reliability standards of each disk drive within the system, using an enhanced implementation of Self-Monitoring Analysis and Reporting Technology (SMART) tools. As typically implemented in the storage industry, SMART tools indicate whether certain thresholds have been exceeded, which gives an alert that a disk is at risk for failure and therefore needs to be replaced.

However, as implemented in XIV Storage System, the SMART diagnostic tools, coupled with intelligent analysis and low tolerance thresholds, provide an even greater level of refinement of the disk behavior diagnostic tests and the performance and reliability driven reaction. For example, the XIV Storage System measures the specific values of parameters including, but not limited to these possibilities:

- *Reallocated sector count*: If the disk encounters a read or write verification error, it designates the affected sector as “reallocated” and relocates the data to a reserved area of spare space on the disk. This spare space is a parameter of the drive and is *not* related in any way to the system reserve spare capacity that is described in “Global spare capacity” on page 34.
- *Disk temperature*: The disk temperature is a critical factor that contributes to premature drive failure and is constantly monitored by the system.
- *Raw read error*: The raw read error count provides an indication of the condition of the magnetic surface of the disk platters and is carefully monitored by the system to ensure the integrity of the magnetic media itself.
- *Spin-up time*: The spin-up time is a measure of the average time that is required for a spindle to accelerate from zero to 7200 rpm. The XIV Storage System recognizes abnormal spin-up time as a potential indicator of an impending mechanical failure.

Likewise, for additional early warning signs, the XIV Storage System continually monitors other aspects of disk-initiated behavior, such as spontaneous reset or unusually long latencies. The system intelligently analyzes this information to reach crucial decisions concerning disk deactivation and phase out. The parameters involved in these decisions allow for a sensitive analysis of the disk health and performance.

Redundancy-supported reaction

The XIV Storage System incorporates *redundancy-supported reaction*, which is the provision that uses the distributed redundant data scheme by intelligently redirecting reads to the secondary copies of data. This extends the system’s tolerance of above-average disk service time when accessing primary data locations. The system reinstates reads from the primary data copy when the transient degradation of the disk service time has subsided. A redundancy-supported reaction might be triggered by an underlying potential disk error that is ultimately managed autonomically by the system according to the severity of the exposure, as determined by ongoing disk monitoring.

Flexible handling of dirty data

In a similar manner to the redundancy-supported reaction for read activity, the XIV Storage System can also make convenient use of its redundant architecture to consistently maintain write performance. Because intensive write activity directed to any particular volume is distributed across all modules and drives in the system, and the cache is independently managed within each module, the system is able to tolerate sustained write activity to an under-performing drive by effectively maintaining a considerable amount of “dirty,” or

unwritten, data in cache. This situation potentially circumvents any performance degradation resulting from the transient and anomalous service time of a particular disk drive.

Non-Disruptive Code Load

Non-Disruptive Code Load (NDCL) enables upgrades to the XIV Storage System Software from a current version (starting with Version 11.1.0) to a later version without disrupting user applications.

The code upgrade is run on all modules in parallel and the process is fast enough to have no impact on host applications.

No data migration or rebuild process is allowed during the upgrade. Mirroring, if any, is suspended during the upgrade and automatically reactivated upon completion.

Storage management operations are also not allowed during the upgrade, although the status of the system and upgrade progress can be queried. It is also possible to cancel the upgrade process up to a point of no return.

The NDCL does not apply to specific component's firmware upgrades (for example, module basic input/output system (BIOS) and host bus adapter (HBA) firmware). These components require a phase in/phase out process of the impacted modules.

Note: Starting from Release 11.3, hot firmware upgrade is supported for certain Unified System Management (USM), SAS disk, and flash cache components. For details, contact your IBM technical support.

Optional flash cache upgrade

Flash cache can be upgraded by using the Miscellaneous Equipment Specification (MES) process if the XIV system was originally installed without flash cache.

Flash cache installation is a concurrent, nondisruptive process to be performed by an IBM technician.

After flash cache is inserted into the XIV modules and enabled, it is immediately ready for use by the XIV software. Depending on the use profile, it can help improve performance.

3

IBM XIV architecture, components, and planning

This chapter describes the hardware architecture of the IBM XIV Storage System Gen3 family. It covers the physical components for the newer Model 214 and the Model 114, including the system rack, interface modules, data modules, management modules, disk drives, network switches, and power distribution devices. It also provides an overview of planning and other tasks required before and after deployment of an IBM XIV Storage System

There are two machine types associated with the XIV Storage System: the 2810 and the 2812. Both machine types have the standard warranty period. IBM XIV systems with a machine type 2810 have a one-year warranty and ones with a machine type of 2812 have three-year warranties.

This chapter includes the following topics:

- IBM XIV Storage System Gen3
- Hardware planning overview

© Copyright IBM Corp. 2014. All rights reserved.

3.1 IBM XIV Storage System Gen3

The XIV Storage System Gen3 Model 214 shares all of the concepts and software architecture of its predecessor Model 114. Compared to the Model 114, the Model 214 consists of hardware upgrades that deliver significant performance improvements. The primary hardware component upgrades are in the following areas:

- Interface and data modules
  - Processor
  - Memory
  - Power Supply Unit
  - 1 TB[1], 2 TB, 3 TB, 4 TB, and 6 TB serial-attached SCSI (SAS)-self-encrypting drives (SEDs)
  - 400 GB flash drive cache option or 800 GB flash drive with 4 TB configurations
- Interface modules
  - Two-port 10 GbE adapter

The XIV Gen3, shown in Figure 3-1, is a scalable enterprise storage system based on a grid array of hardware components. The architecture offers the highest performance through maximized and balanced use of all disks, a true distributed cache implementation that produces exceptional performance. It also offers superior reliability through its distributed architecture, redundant components, self-monitoring, and self-healing attributes.

Figure 3-1 IBM XIV Storage System Gen3: Front and rear views

[1] The 1 TB drives are actually stripped-down 2 TB drives.

The XIV Storage System Gen3 can use flash drives, which are also referred to as solid-state drives (SSDs), in each module. Flash drives are used as secondary read cache between disks and the dynamic random access memory (DRAM) cache. Data is always placed in both cache units, and disk access is avoided when data is in SSD/flash drive cache. Because the SSD/flash drive cache is multiple times larger than the DRAM cache, the flash cache is highly effective in regard to the read cache hit ratio.

Table 3-1 shows the main hardware components.

Table 3-1 XIV Storage System Gen3 hardware overview

Component	Model 114	Model 214 (1 GbE)	Model 214 (10 GbE)
Rack	IBM T42 42U	IBM T42 42U	IBM T42 42U
All Modules	Intel Quad core CPU 24 GB DDR memory 12x SAS drives 2x Redundant Power Supplies/module 2U	Intel Six core CPU 48 GB DDR3 1.3 GHz memory 12x SAS drives 2x High Efficient. Power Supplies 2U	Intel Six core CPU 48 GB DDR3 1.3 GHz memory 12x SAS drives 2x High Efficient. Power Supplies 2U
Interface Module	In addition to data module: - 2x iSCSI Port on Module 4 - 4x iSCSI Ports on active Modules - 4x 8 Gbps FC ports	In addition to data module: - 2x iSCSI Port on Module 4 - 4x iSCSI Ports on active Modules - 4x 8 Gbps FC ports	In addition to data module: - 2x iSCSI Ports 10 GbE - 4x 8 Gbps FC ports
Switch	- 2x 36 ports InfiniBand	- 2x 36 ports InfiniBand	- 2x 36 ports InfiniBand
Maintenance Module and Modem	- 1U Maintenance Module - 1x Modem	- 1U Maintenance Module - 1x Modem	- 1U Maintenance Module - 1x Modem
UPS	- 3x UPS	- 3x UPS	- 3x UPS
ATS	- 1x ATS	- 1x ATS	- 1x ATS
Full Rack	- 15x Modules - 180x disks - 2 TB disks = 161 TB - 3 TB disks = 243 TB - 360 GB memory (cache) - 6 TB flash cache - 6x interface modules - 24x FC ports - 22x iSCSI ports - 9x data modules	- 15x Modules - 180x disks - 2 TB disks = 161 TB - 3 TB disks = 243 TB - 4 TB disks = 325 TB - 6 TB disks = 489 TB - 360 GB cache with 2 TB, 3 TB disks - 720 GB cache with 4 TB, 6 TB disks - 6 TB flash cache - 12 TB flash cache with 4 TB, 6 TB disks - 6x interface modules - 24x FC ports - 22x iSCSI ports - 9x data modules	- 15x Modules - 180x disks - 2 TB disks = 161 TB - 3 TB disks = 243 TB - 4 TB disks = 325 TB - 6 TB disks = 489 TB - 360 GB cache with 2 TB, 3 TB disks - 720 GB cache with 4 TB, 6 TB disks - 6 TB flash cache - 12 TB flash cache with 4 TB, 6TB disks - 6x interface modules - 24x FC ports - 12x iSCSI ports - 9x data modules

All XIV Gen3 hardware components are delivered preinstalled in a standard IBM T42 19-inch rack. Data and interface modules provide the processing, caching, and storing of data. All modules can be considered data modules in that they each contain a processor, memory, and 12 serial-attached SCSI (SAS) drives. The SAS drives can be 2 TB and 3 TB. The SAS-SED drives can be 2 TB, 3 TB, 4 TB, or 6 TB. The interface modules are data modules but with more capabilities. The interface modules have unique software and hardware that provide Fibre Channel and IP network Small Computer System Interface (iSCSI) host connectivity.

Important: Disks of different sizes cannot be intermixed within the same XIV Storage System configuration. XIV Storage System Software Version 11.0.1 or later is required for using 3 TB disks. Also, XIV Storage System Software Version 11.4.0 or later is required for using the SAS-SED drives in 2 TB, 3 TB, or 4 TB and XIV Storage System software Version 11.5.1 or later is required for using the SAS-SED drives in 6 TB.

The bottom of the rack contains three uninterruptible power supplies (UPSes) that supply power to all XIV Storage System components. They also provide enough battery backup power for emergency shutdown if main power is lost for more than 30 seconds.

A maintenance module and two InfiniBand switches are installed near the middle of the rack, just above module 6. The InfiniBand network provides redundant and fast communication paths between the modules. This grid network ensures communication between all modules even if one of the switches or a cable connection fails. It also provides the capabilities for parallelism and the execution of a data distribution algorithm that contributes to the performance of the XIV Storage System.

All cabling between the modules and switches, including the internal power connections, are fully redundant using two sets of cables. All cables use industry standard plugs. The XIV rack comers precabled for all 15 modules even if smaller configuration is ordered.

3.1.1 Fully populated configurations

A fully populated rack contains 9 data modules and 6 interface modules, for a total of 15 modules. The total usable capacity is 161 TB for a complete system when equipped with 2 TB drives, 243 TB of usable capacity when fully populated with 3 TB drives, 325 TB of usable capacity when fully populated with 4 TB drives, and 489 TB of usable capacity when fully populated with 6 TB drives.

Because each module in Model 214 contains up to 48 GB of memory (in the 4 TB disk configuration), a full rack contains up to 720 GB of memory that can be used to handle host read and write I/O requests (Model 114 has up to 24 GB installed per module, which equates to 360 GB for the full rack). The SSD extended caching option adds 400 GB of read cache capacity to each module, for a total of 6 TB in a fully populated configuration (15 modules). In the 4 TB or 6 TB version of the Model 214, the flash cache option can be ordered in 800 GB of read cache capacity to each module, for a total of 12 TB in a fully populated configuration (15 modules).

Summary:

- Each interface module of the XIV Gen3 Model 114 contains four 8 Gbps Fibre Channel ports and one 4-port 1 GbE adapter, except interface module 4, which has only two iSCSI ports.
- Each interface module of the XIV Gen3 Model 214 (1 GbE) contains four 8 Gbps Fibre Channel ports and one 4-port 1 GbE adapter, except interface module 4, which has only two iSCSI ports.
- Each interface module of the XIV Gen3 Model 214 (10 GbE) contains four 8 Gbps Fibre Channel ports and one 2-port 10 GbE adapter.

3.1.2 Partially populated configurations

The XIV Storage System Gen3 is also available in partially populated configurations of 6, 9, 10, 11, 12, 13, or 14 modules, allowing for more granularity of capacity options. Modules can be added as capacity needs increase up to 15 modules.

Different size drives cannot be intermixed within the same IBM XIV Storage System.

The SSD Caching extension is also available with partially populated configurations. The SSD extension is required for each module present in the partial configuration.

In Figure 3-2, certain interface modules are labeled *Disabled*. This label means that the interface module is not running the special software that characterizes an interface module and the host interface adapters are not functional. They still function as data modules. As modules are added to the XIV Storage System configurations, these interface modules become *Enabled* from an interface perspective.

Certain partial rack configurations do not use all host attachment interface ports even though they might be physically present. The interface ports are activated automatically as more modules are added to the system.

Details about these configuration options and the various capacities, drives, ports, and memory are provided in Figure 3-2.

Rack Configuration									
Total number of modules (Configuration type)		6 partial	9 partial	10 partial	11 partial	12 partial	13 partial	14 partial	15 full
Total number of data modules		3	3	4	5	6	7	8	9
Total number of interface modules		3	6	6	6	6	6	6	6
Number of active interface modules		2	4	4	5	5	6	6	6
Interface module 9 state			Disabled	Disabled	Enabled	Enabled	Enabled	Enabled	Enabled
Interface module 8 state			Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled
Interface module 7 state			Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled
Interface module 6 state		Disabled	Disabled	Disabled	Disabled	Disabled	Enabled	Enabled	Enabled
Interface module 5 state		Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled
Interface module 4 state		Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled
FC ports		8	16	16	20	20	24	24	24
iSCSI ports (1 Gbps – mod 114)		6	14	14	18	18	22	22	22
iSCSI ports (10 Gbps – mod 214)		4	8	8	10	10	12	12	12
Number of disks		72	108	120	132	144	156	168	180
Usable capacity (1 / 2 / 3 / 4 / 6 TB)		28 TB 55 TB 84 TB 112 TB 169 TB	44 TB 88 TB 132 TB 177 TB 267 TB	51 TB 102 TB 154 TB 207 TB 311 TB	56 TB 111 TB 168 TB 225 TB 338 TB	63 TB 125 TB 190 TB 254 TB 382 TB	67 TB 134 TB 203 TB 272 TB 409 TB	75 TB 149 TB 225 TB 301 TB 453 TB	81 TB 161 TB 243 TB 325 TB 489 TB
# of CPUs (one per Module)		6	9	10	11	12	13	14	15
Memory (24 GB per module w 1/2/3 TB) Memory (48 GB per module w 4/6 TB)		144 GB 288 GB	216 GB 432 GB	240 GB 480 GB	264 GB 528 GB	288 GB 576 GB	312 GB 624 GB	336 GB 672 GB	360 GB 720 GB
{Optional for 1, 2 ,3, 4, 6 TB XIVs} 400 GB Flash Cache {Optional for 4, 6 TB XIVs} 800 GB Flash Cache		2.4 TB 4.8 TB	3.6 TB 7.2 TB	4.0 TB 8.0 TB	4.4 TB 8.8 TB	4.8 TB 9.2 TB	5.2 TB 10.4 TB	5.6 TB 11.2 TB	6.0 TB 12.0 TB
Power (kVA) - Model 281x-214 / with SSD		2.5 / 2.6	3.6 / 3.9	4.0 / 4.3	4.3 / 4.6	4.7 / 5.09	5.0 / 5.4	5.5 / 5 .8	5.8 / 6.2

Figure 3-2 Partial configurations Gen3 - Model 214

For more information about *usable capacity*, see 2.3, “Full storage virtualization” on page 24.

3.1.3 Capacity on Demand

XIV Storage System Gen3 has an optional pricing model that allows clients to purchase data storage capacity as it is needed. This feature is called *Capacity on Demand* (*CoD*).

Restriction: Capacity on Demand is not available for the 1 TB drive option.

CoD means that an XIV Storage System Gen3 can be ordered with a certain amount of authorized storage capacity, as well as extra storage capacity that is not intended to be used initially and will be purchased later as this additional storage is allocated to XIV Storage System storage pools.

IBM is alerted that this storage is being used and an invoice is generated to authorize the use of this additional storage. This situation is referred to as purchasing a *module activation feature*. The XIV sends reports to IBM about allocated storage through the Call Home feature (modem or email), and IBM invoices clients if licensed capacity is exceeded. An XIV Storage System that is in the CoD program must be able to call home to IBM. (This is mandatory.)

The advantage of CoD is that storage purchases can be deferred until the storage capacity is needed. There is a small price premium associated with purchasing CoD storage over the traditional way of paying for all storage capacity in the whole configuration up front. It is common for clients to stay in the XIV Capacity on Demand program until the XIV Storage System configuration has reached nine modules and then switch to purchasing normal (non-CoD) modules to further expand the capacity of the XIV Storage System. Switching from a CoD configuration to a non-CoD configuration is allowed, but then you cannot later purchase more CoD capacity.

The most important aspect of all CoD configurations is that the system is delivered with the same hardware configuration as the non-CoD system. All modules are active. The XIV uses the performance capabilities of all the disk drives in the configurations, purchased storage, and CoD de-allocated and unpurchased storage.

Benefit: All CoD configurations are delivered with the same hardware as a full system, and all modules and all ports are active. The restriction is only on the capacity that you can use.

This scenario means that data is initially distributed over all of the drives in the XIV Storage System configuration. There are no restrictions on the normal functions. The full performance capacity of all modules in the rack is used, even though some of the modules are unused CoD modules.

The basic configuration rule to keep in mind about the XIV Storage System Gen3 CoD program is that there can be only up to three de-allocated CoD modules in the configuration. This configuration means that every valid CoD configuration always has one, two, or three de-allocated CoD modules in the configuration.

There are several considerations for Model 114 and Model 214 CoD configurations:

- The minimum CoD system configuration is six modules. This configuration requires a minimum of three and a maximum of five CoD activation features (between three and five modules are already purchased and between one and three modules are available for purchase).

- The next valid CoD configuration is nine modules. This configuration requires a minimum of six and a maximum of eight CoD activation features (between six and eight modules are already purchased and between one and three modules are available for purchase).
- Participation in the CoD program is valid if there are unactivated CoD modules in the system.
- When activating (allocating) CoD modules, ensure that there are remaining unactivated CoD modules in the configuration to continue your participation in the CoD program.
- Any order that fully activates a CoD configuration and does not include additional unactivated CoD features indicates the client's desire to exit the CoD program.
- There are explicit CoD term limits:
  - One-year CoD term: At the end of one year, the system exits the CoD program and automatically activates all remaining unactivated (and de-allocated) CoD modules, generating client invoices for all of the remaining CoD storage capacity.
  - Two-year CoD term: At the end of two years, the system exits the CoD program and automatically activates all remaining unactivated (and de-allocated) CoD modules, generating client invoices for all of the remaining CoD storage capacity.
  - Three-year CoD term: At the end of three years, the system exits the CoD program and automatically activates all remaining unactivated (and de-allocated) CoD modules, generating client invoices for all of the remaining CoD storage capacity.
- A warranty for all physically installed modules begins when they are initially powered on. They are being used to store data and contributing to performance regardless of whether they have been purchased or not.
- Allocating more CoD modules effectively creates a rolling CoD term.
- XIV Storage System Call Home capabilities must be enabled to participate in the XIV CoD program.

Figure 3-3 on page 68 shows an example of the valid capacities in decimal TB of an XIV Storage System Gen3 with 2 TB drives. The table shows that the capacity per CoD activation varies depending on the exact configuration.

For example, a machine with 10 physical modules has 102.6 TB of usable capacity. If the client purchased seven activations, they can use 71.82 TB of that 102.6 TB. With each extra activation, the client could use an extra 10.26 TB of usable capacity. If they purchase a total of 10 activations without purchasing any extra physical modules, they exit the CoD program and are able to use all of the 102.6 TB. If they instead purchase one more activation (for a total of eight) and one more physical module (for a total of 11), they are able to use 81.091 TB of 111.5 TB of usable capacity. Each extra activation now buys 10.136 TB.

Figure 3-3 shows an example of the valid capacities in TBs of an XIV Storage System Gen3 with 2 TB drives.

2TB	6 Physical Modules	9 Physical Modules	10 Physical Modules	11 Physical Modules	12 Physical Modules	13 Physical Modules	14 Physical Modules	15 Physical Modules
Maximum Capacity -->	**55700**	**88000**	**102600**	**111500**	**125900**	**134900**	**149300**	**161300**
Capacity per CoD Activation -->	**9283**	**9778**	**10260**	**10136**	**10492**	**10377**	**10664**	**10753**
3 CoD Activations -->	27850							
4 CoD Activations -->	37133							
5 CoD Activations -->	46417							
6 CoD Activations -->	**55700**	58667						
7 CoD Activations -->		68444	71820					
8 CoD Activations -->		78222	82080	81091				
9 CoD Activations -->		**88000**	92340	91227	94425			
10 CoD Activations -->			**102600**	101364	104917	103769		
11 CoD Activations -->				**111500**	115408	114146	117307	
12 CoD Activations -->					**125900**	124523	127971	129040
13 CoD Activations -->						**134900**	138636	139793
14 CoD Activations -->							**149300**	150547
15 CoD Activations -->								**161300**

Figure 3-3 CoD capacities for an XIV Storage System Gen3 with 2 TB drives

Figure 3-4 shows an example of the valid capacities in TBs of an XIV Storage System Gen3 with 3 TB drives.

3TB	6 Physical Modules	9 Physical Modules	10 Physical Modules	11 Physical Modules	12 Physical Modules	13 Physical Modules	14 Physical Modules	15 Physical Modules
Maximum Capacity -->	**84100**	**132800**	**154900**	**168300**	**190000**	**203600**	**225300**	**243300**
Capacity per CoD Activation -->	**14017**	**14756**	**15490**	**15300**	**15833**	**15662**	**16093**	**16220**
3 CoD Activations -->	42050							
4 CoD Activations -->	56067							
5 CoD Activations -->	70083							
6 CoD Activations -->	**84100**	88533						
7 CoD Activations -->		103289	108430					
8 CoD Activations -->		118044	123920	122400				
9 CoD Activations -->		**132800**	139410	137700	142500			
10 CoD Activations -->			**154900**	153000	158333	156615		
11 CoD Activations -->				**168300**	174167	172277	177021	
12 CoD Activations -->					**190000**	187938	193114	194640
13 CoD Activations -->						**203600**	209207	210860
14 CoD Activations -->							**225300**	227080
15 CoD Activations -->								**243300**

Figure 3-4 CoD capacities for an XIV Storage System Gen3 with 3 TB drives

Figure 3-5 shows an example of the valid capacities in TBs of an XIV Storage System Gen3 with 4 TB drives.

4TB	6 Physical Modules	9 Physical Modules	10 Physical Modules	11 Physical Modules	12 Physical Modules	13 Physical Modules	14 Physical Modules	15 Physical Modules
Maximum Capacity -->	**112500**	**177700**	**207100**	**225100**	**254100**	**272300**	**301300**	**325400**
Capacity per CoD Activation -->	**18750**	**19744**	**20710**	**20464**	**21175**	**20946**	**21521**	**21693**
3 CoD Activations -->	56250							
4 CoD Activations -->	75000							
5 CoD Activations -->	93750							
6 CoD Activations -->	**112500**	118467						
7 CoD Activations -->		138211	144970					
8 CoD Activations -->		157956	165680	163709				
9 CoD Activations -->		**177700**	186390	184173	190575			
10 CoD Activations -->			**207100**	204636	211750	209462		
11 CoD Activations -->				**225100**	232925	230408	236736	
12 CoD Activations -->					**254100**	251354	258257	260320
13 CoD Activations -->						**272300**	279779	282013
14 CoD Activations -->							**301300**	303707
15 CoD Activations -->								**325400**

Figure 3-5 CoD capacities for an XIV Storage System Gen3 with 4 TB drives

Figure 3-6 shows an example of the valid capacities in TBs of an XIV Storage System Gen3 with 6 TB drives.

6TB	6 Physical Modules	9 Physical Modules	10 Physical Modules	11 Physical Modules	12 Physical Modules	13 Physical Modules	14 Physical Modules	15 Physical Modules
Maximum Capacity -->	**169300**	**267400**	**311600**	**338700**	**382300**	**409700**	**453300**	**489500**
Capacity per CoD Activation -->	**28217**	**29711**	**31160**	**30791**	**31858**	**31515**	**32379**	**32633**
3 CoD Activations -->	84650							
4 CoD Activations -->	112867							
5 CoD Activations -->	141083							
6 CoD Activations -->	**169300**	178267						
7 CoD Activations -->		207978	218120					
8 CoD Activations -->		237689	249280	246327				
9 CoD Activations -->		**267400**	280440	277118	286725			
10 CoD Activations -->			**311600**	307909	318583	315154		
11 CoD Activations -->				**338700**	350442	346669	356164	
12 CoD Activations -->					**382300**	378185	388543	391600
13 CoD Activations -->						**409700**	420921	424233
14 CoD Activations -->							**453300**	456867
15 CoD Activations -->								**489500**

Figure 3-6 CoD capacities for an XIV Storage System Gen3 with 6 TB drives

3.1.4 Advanced System Placement

IBM Advanced System Placement (ASP) is a pay-as-you-go program that lets organizations purchase IBM XIV Storage Systems for only a fraction of the price upon installation. The ASP program complements the Capacity on Demand pricing program.

For details and official terms of the offer, see the IBM Announcement Letter Number 314-041 about IBM System Storage Advanced System Placement:

http://ibm.co/ZF5O5J

This acquisition program serves a highly flexible way to provide XIV storage. It is available for new and current IBM clients and is ideal for cloud service providers. It is also a good fit for clients with rapidly growing storage requirements who need highly flexible ways to add storage.

Advanced System Placement might be good to consider for clients who meet one or more the following situations:

- Clients who buy more than 160 TB of usable capacity per year
- Clients with unpredictable growth needs who are not sure when they will need the extra capacity and would like to defer payment
- Clients about to launch a major new application
- Clients who are acquiring another company
- Clients who are deploying private clouds
- Client who expect a new contract or client

The acquisition is simple. IBM installs a full system with partial payment, based on initial use. When use and capacity requirements grow, there is no need for hardware configuration changes or intervention from an IBM sales support representative nor any key activation. In contrast to the CoD program, Advanced System Placement does not require detailed capacity monitoring. It also provides the ability to run several parallel ASP tracks for multi-site data centers and client needs. Clients sign a contract for standard a one-year fixed term.

A financed version of the ASP program is available through IBM Global Financing. By combining this program with a hardware lease program, payments become predictable and manageable. The offering can be proposed as a capital acquisition model or, in many cases, as an operating expense (Opex model), with payments spread over 36 months (60% stepping up to 100% or no payments stepping up to 100%).

For each agreement, there may be one initial machine and any number of additional machines called “subsequent machines” that are acquired one at a time. There is no limit to the number of agreements that the client may have in place at any one time.

Figure 3-7 on page 71 depicts a one-year acquisition model. IBM installs a fully populated 15-module XIV system with partial payment (60% of total price) based on initial use. The client gets full XIV performance from the first day. Depending on client growth rates and service level agreement (SLA), IBM procures subsequent systems in advance of client needs, typically at 70% use.

IBM charges one US dollar for initial invoices for the second system and the remaining 40% charges for the first machine. In the same manner, IBM will deliver one additional system whenever the previous machine reaches the expansion threshold (70% use, in this example) and charges the one US dollar initially for the additional system. At that time, the client must

also make the remaining payment (that is the 100% price less $1) for the previous system. In any case, at the end of the one-year term, the client pays for any delivered system.

Important: Remember that with ASP, the capacity increments are full systems: 161 TB, 243 TB, 325 TB, or 489 TB.

The program supports the XIV Storage System machine type 2810 or 2812 and model 214 and later.

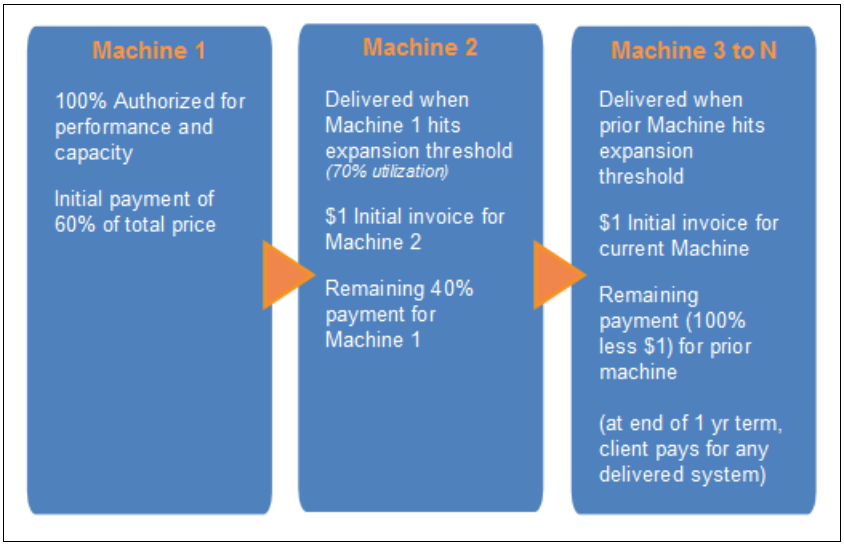

Machine 1
100% Authorized for performance and capacity
Initial payment of 60% of total price
Machine 2
Delivered when Machine 1 hits expansion threshold
(70% utilization)
$1 Initial invoice for Machine 2
Remaining 40% payment for Machine 1
Machine 3 to N
Delivered when prior Machine hits expansion threshold
$1 Initial invoice for current Machine
Remaining payment (100% less $1) for prior machine
(at end of 1 yr term, client pays for any delivered system)

Figure 3-7 Advanced System Placement pricing model

The client must sign two documents:

- Advanced System Placement Agreement Z126-6245-US or CA
  - The Agreement contains the terms of the program
  - The Agreement is signed one time for each client
- Advanced System Placement Supplement Z126-6246-US or CA
  - The Supplement contains the details of the machine and time frame
  - A Supplement must be signed for each machine

One *Agreement* consists of an initial machine and any number of subsequent machines. There is no limit to the number of subsequent machines that a client may have under an agreement, but the agreements must be acquired one at a time.

Table 3-2 compares the IBM XIV Storage System CoD to ASP program.

Table 3-2 XIV CoD and ASP comparison

Attribute	XIV CoD	XIV ASP
Client pays for capacity used	By XIV module increment: ~10, 16, or 22 TB. Consumption determined last day of each month.	Buys full XIV system, payment split into 2 parts. Second invoice is either when use > 70% or at the end of the period.
Minimum capacity purchased	Initial system as small as 27 TB, increments of XIV activated modules (10, 16, 22 TB). Standard config is 3 extra modules (can RPQ more).	Full XIV systems only: 161 TB, 243 TB, 325 TB, or 489 TB.
Term period	1 - 3 years	1 year, standard
Acquired	Purchase or lease XIV system with CoD modules under CoD contract.	Purchase or lease full XIV system under ASP contract.
Client charges	Point in time: a) Initial system purchase b) when capacity used exceeds authorized, or c) at end of CoD period.	When system is shipped: If 1st system: 60%. Otherwise: $1. When use > 70%, remainder is due.
Sold through	IBM direct and Business Partner sales, purchase or lease (IGF)	IBM direct and Business Partner sales (IGF lease only)

3.1.5 XIV Cloud Storage for Service Providers

IBM XIV Cloud Storage for Service Providers offers feature options and innovative pay-per-need flexibility to empower cloud providers to cater to tenant requirements with exceptional efficiency. With this offering, cloud service providers and managed service providers can choose XIV functions based on the needs and cost sensitivities of their clients so that they can pay for what they need when they need it.

XIV Cloud Storage for Service Providers features are optimized for cloud services and offer the following advanced function licenses for purchase as needed (see Figure 3-8 on page 73):

- Efficiency: Flash caching
- Elasticity: HyperScale mobility
- Security: Encryption
- Continuity: Mirroring

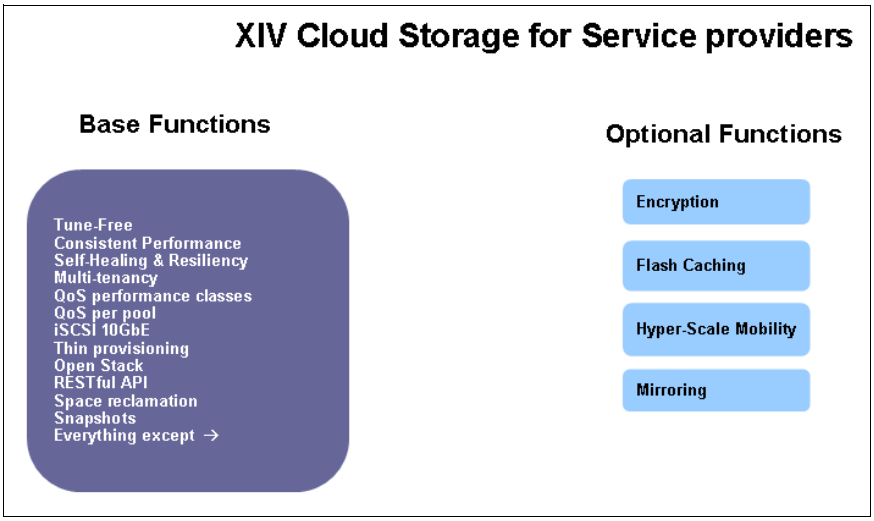


Figure 3-8 XIV Cloud Storage for Service Providers, base and optional functions

Clients with fewer than three optional features may add features. Clients with three optional features will automatically receive the full set of components, after which there will be no need to order other optional components.

XIV Cloud Storage for Service Providers license requires one IBM XIV Storage System Gen3 machine type 2810-214 or 2812-214 (full systems: 6 interface modules, 9 data modules, and feature number 0815 are required) in any of the validly recordable configurations.

3.1.6 XIV Storage System Model 114 and Model 214 hardware components

The system architecture of the XIV Storage System is designed, wherever possible, to use off-the-shelf components (except for the Automatic Transfer Switch (ATS)) that are not dependent upon specifically designed hardware or proprietary technology. This architecture is optimized for flexibility so that as newer and higher performing components are made available in the marketplace, development is able to incorporate this newer technology into the base system design at a faster pace than was traditionally possible. The following sections describe the hardware components that build up the XIV Storage System.

For more detailed planning information, see 3.2, “Hardware planning overview” on page 94. Also see the *IBM XIV Storage System Planning Guide.* The latest version is available in both PDF format and HTTP format in the IBM XIV Storage System documentation in the IBM Knowledge Center (IBM XIV Storage System > IBM XIV Gen3 281x-11x and 281x-21x > Planning):

`http://publib.boulder.ibm.com/infocenter/ibmxiv/r2/index.jsp`

All XIV Storage System Gen3 configurations include the following components:

- Rack
- Power components
- Data modules and interface modules
- InfiniBand module interconnect
- Patch panel
- Support hardware

Hardware: Because the XIV Storage System hardware architecture is designed using off-the-shelf components and can easily adapt to new hardware, it is possible that there will be hardware changes before this book is refreshed.

3.1.7 Rack

All of the XIV Storage System hardware components are installed in an IBM T42 rack, as shown in Figure 3-9. Adequate space is provided to house all components and to properly route all cables. The rack door and side panels can be locked with a key to prevent unauthorized access to the installed components. This XIV Gen3 rack security kit is available by ordering RPQ 8S1190.

The XIV Storage System Gen3 rack should be considered a dedicated unit for XIV Storage System. Spare space within the rack cannot be used for other equipment.

Rack space: Unused rack space cannot be used for other purposes.

Figure 3-9 XIV Storage System Gen3 rack

The XIV Storage System Gen3 rack has four optional features, as described in Table 3-3.

Table 3-3 Rack features for XIV Storage System Gen3

Feature code	Feature description
0080	Ruggedized rack option
0082	Rear-door heat exchanger
0200	Weight and height, reduced shipping cost
0203	Radio frequency identification device

Ruggedized rack option

The ruggedized rack is an optional kit for stabilizing the XIV Gen3 rack so that the rack complies with IBM earthquake-resistance standards. It is important for XIV systems to be adequately restrained during earthquakes to prevent human injury and limit potential damage to critical system components, such as hard disk drives. The optional ruggedized rack option includes hardware that secures the rack to the floor and cross braces on the front and rear of the rack, which prevents the rack from twisting. The earthquake-resistance option brings the XIV Storage System into compliance with earthquake-resistance objectives that are documented in *Earthquake Resistance for IBM Hardware Products* (IBM Corporate Bulletin C-B 1-9711-009 9202).

Installation of the required floor hardware and the earthquake resistance kit is disruptive. If the earthquake resistance kit is installed on an existing XIV Storage System, the XIV Storage System must be turned off and temporarily moved while the floor is prepared and the kit is installed.

The rack tie downs are intended for securing a frame weighing 1134 kg (2500 lbs) per rack. These tie downs are designed to secure the rack on either a non-raised floor or a raised floor installation.

IBM has made every effort to conduct limited tests, but not all situations are tested, and the drawings and data are provided on an *as is* basis, with no warranty of any kind, express or implied. Rather, IBM has provided the information to help procure the parts needed. You can either install the design or request IBM to install the design as a service.

For more details, see the *IBM XIV Storage System Gen3 Models 281x-11x and 281x-21x Planning Guide*, SC27-5412.

Rear-door heat exchanger

The rear-door heat exchanger is an optional water-cooled device that is mounted on the rear of an XIV Storage System to cool the air that is heated and exhausted by devices inside the rack. A supply hose delivers chilled water to the heat exchanger. A return hose delivers warm water back to the water pump or chiller (referred to as the secondary cooling loop). The primary cooling loop supplies the building chilled water to secondary cooling loops, air conditioning units, and so on. The hoses for the secondary cooling loop are not included with this option.

The XIV Storage System Gen3 rack that is installed with this option can be on a raised floor or a non-raised floor. The required planning for the rear-door heat exchanger is extensive. Details can be found in the *IBM XIV Storage System Planning Guide*, SC27-5412.

Weight and height reduction

The XIV Gen3 weight and height reduction feature removes the top four modules from the rack and ships them separately. This option leaves the rack weighing approximately 1018 kg (2240 lbs). This option reduces the weight loading on floors and elevators that are used when moving the XIV Storage System rack from the truck to the final position in the installation site. An IBM Service Support Representative installs the components that were shipped separately during system installation.

Also, if this feature is ordered, the IBM SSR can also remove the rack top cover to reduce the rack total height to help with moving the rack through low clearance obstacles.

Power consumption

Table 3-4 lists the power consumption of the XIV Gen3 module configurations. The measurements were taken in an environment with a room temperature of 25° C (77° F).

Table 3-4 XIV power consumption

Modules in rack	# disk drives	Model 281x-114 with SSDs in kVA High performance (1 and 2 TB / 3 TB)	Model 281x-214 with SSDs in kVA High performance (2 TB / 3 TB / 4 TB / 6 TB)
6	72	2.9 / 2.9	2.5 / 2.6 / 2.6 / 2.6
9	108	4.3 / 4.2	3.6 / 3.8 / 3.9 / 3.9
10	120	4.7 / 4.7	4.0 / 4.2 / 4.3 / 4.3
11	132	5.1 / 5.2	4.3 / 4.5 / 4.6 / 4.6
12	144	5.5 / 5.7	4.7 / 4.9 / 5.0 /5.0
13	156	5.9 / 6.1	5.0 / 5.3 / 5.4 / 5.4
14	168	6.5 / 6.7	5.5 / 5.7 / 5.8 / 5.8
15	180	6.9 / 7.2	5.8 / 6.1 / 6.2 / 6.2

Important: Model 214 reduces power consumption by up to 16% compared to previous models and generations, thanks to the high efficiency power supply unit (PSU) and the new six-core CPUs.

Thermal dissipation

Table 3-5 indicates the cooling (thermal dissipation) requirements for a 15-module rack. To support capacity upgrades, the installation site must provide cooling capacity to support full-rack configurations.

Table 3-5 Thermal dissipation in kBTU/hour

Model-Type	2TB	3TB	4TB or 6TB
281x-114	22.7	24.0	N/A
281x-214	19.1	20.1	20.5

Physical space requirements

The physical requirements for the room where the XIV Storage System will be installed must be checked before the arrival of the machine. IBM assists with this activity.

For the latest and most accurate information, see the *IBM XIV Storage System Gen3 Models 281x-11x and 281x-21x Planning Guide*, SC27-5412*:*

- The floor must be able to withstand the weight of a fully configured XIV Storage System Gen3, which is 1040.8kg (2295 lb.). For racks with fewer than 15 modules, subtract 28.6 kg (63 lb.) for each module less than 15 to get the approximate weight requirement for your rack.
- Adequate cooling is required for the XIV Storage System Gen3 configuration ordered. Enough clearance around the system must be left for cooling and service. The thermal dissipation of a fully configured XIV Storage System Gen3 Model 114 with 2 TB drives is 22.7 kBTU/hour and 24 kBTU/hour with 3 TB drives. The thermal dissipation of a fully configured XIV Storage System Gen3 Model 214 with 2 TB drives is 19 kBTU/hour, 20.1 kBTU/hour with 3 TB drives, 20.5 kBTU/hour with 4 TB drives and 6 TB drives. Airflow is from front to back.
- Building features, such as any ramps, elevators, and floor characteristics, must also be considered.

The following measurements in Figure 3-10 are provided for your convenience.

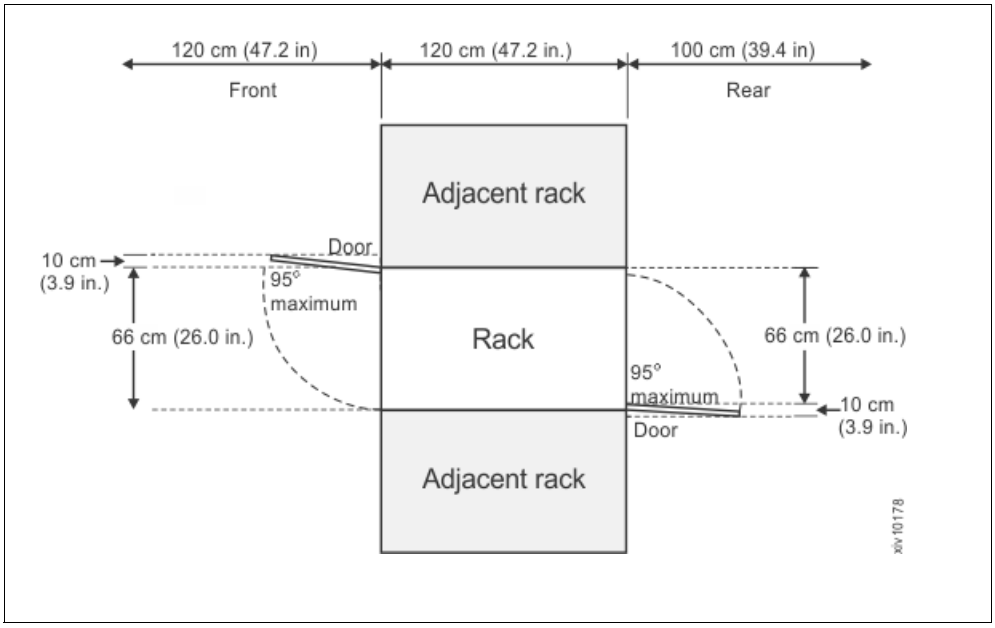


Figure 3-10 Rack clearances: Model 214 and 114

3.1.8 Power components

This section describes power components and their redundancy.

Power redundancy

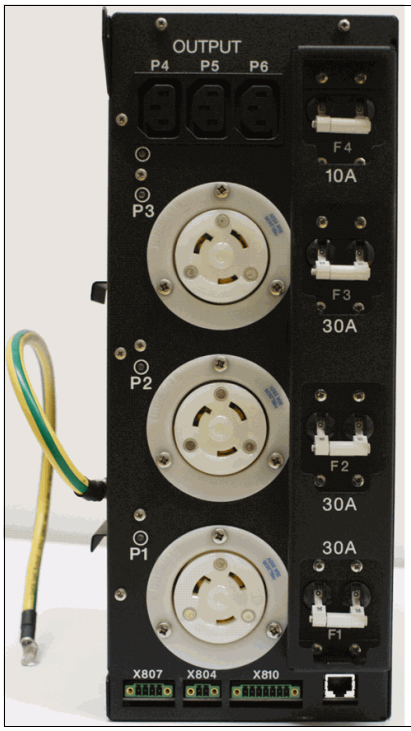
To prevent the complete rack or single components from failing because of power problems, all power components in the XIV Storage System are redundant:

- To protect against loss of client utility power, the XIV Storage System attaches to two independent power sources. The XIV includes two power cables that must be plugged into redundant external power sources. The other ends of the redundant cables are routed inside the XIV rack and attached to the XIV Automatic Transfer Switch (ATS). For most XIV power solutions, the ATS allows power from one of the redundant feeds to power the XIV rack. If there is an external power failure, the ATS transfers the load to the redundant power cord.
- The XIV Storage System has three internal UPSes. It can run on two of these UPSes so that no system component loses power if one of the three UPSes experiences a hardware failure. This redundant design is n+1 because the XIV has one redundant UPS.
- The UPS complex has enough battery reserves to sustain two emergency power shutdowns.
- Each module has two independent power supplies. During normal operation, both power supplies operate on half of the maximal load. If one power supply fails, the remaining operational power supply can take over, and the module continues its operation without any noticeable impact. After the failing power supply is replaced, the power load balancing is restored.

Automatic Transfer Switch

The Automatic Transfer Switch (ATS) shown in Figure 3-11 on page 79 supplies power to all three UPSes and to the maintenance module. If there is a power problem on one line, the ATS reorganizes the power and switches to the other line. The operational components take over the load from the failing power source or power supply. This rearrangement is performed by the ATS in a seamless manner so that the system operation continues without any application impact.

Important: When connecting the XIV Storage System to a 30 A power source, ensure that the facility circuit breakers can handle the inrush currents and differing power loads. For example, D-Curve type breakers provide higher tolerances to allow temporary loads to flow without tripping the facility circuit. Locations in Europe, Middle East, and Africa (EMEA), Australia, and New Zealand should take special note of this information.



Port overview	
P1,P2,P3	UPS power sockets
P4	maintenance module power socket
F1,F2,F3	UPS circuit breaker
F4	maintenance module circuit breaker
X807	ATS monitoring to UPS-3
network port	ATS monitoring to Module-4
P5,P6,X804,X810	not in use
J1,J2	external power

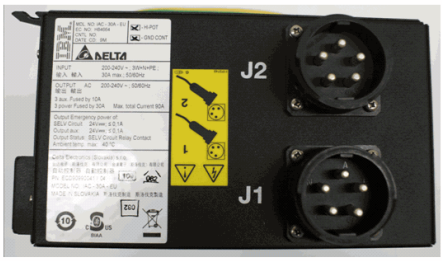


Figure 3-11 Automatic Transfer Switch

The ATS is monitored by the system and generates system event messages in case of problems. The status of the ATS can also be viewed with the **ats_list** XIV Storage System Command-Line Interface (XCLI) command, as shown in Figure 3-12.

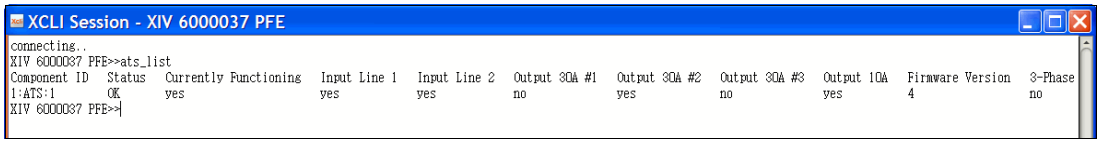

XCLI Session - XIV 6000037 PFE
connecting..
XIV 6000037 PFE>>ats_list
Component ID Status Currently Functioning Input Line 1 Input Line 2 Output 30A #1 Output 30A #2 Output 30A #3 Output 10A Firmware Version 3-Phase
1:ATS:1 OK yes yes yes no yes no yes 4 no
XIV 6000037 PFE>>

Figure 3-12 Output of ats_list command

Four ATS features can be ordered with the XIV Storage System. Order them to correspond to the type of alternating current electric power generation, transmission, and distribution used in the client location. When ordering the XIV Storage System, order the correct feature that meets both local power regulations and client requirements.

The four ATS features are shown in Figure 3-13.

Automatic transfer switch feature code	Description	Feature codes for associated line cords
9890	ATS for worldwide use; single-phase, 60 A	9811, 9813
9891	ATS for United States, Canada, Latin America, and Japan; three-phase delta, 60 A	9820
9892	ATS for EMEA and Asia/Pacific (except Japan); three-phase wye, 30 A	9822
9894	ATS for United States, Canada, Latin America, and Japan; three-phase delta, 30 A	9824

Figure 3-13 ATS features

Figure 3-13 on page 79 also shows the power cords (line cords) that are associated with each ATS. Figure 3-14 shows the details associated with each power cord feature. These power cords are also country-specific. Because each country has specific testing and certification requirements, these power cords cannot be used in countries other than those countries in which they were intended to be used.

Line cord feature code	ATS feature code	Description	Connector (See Note 1.)	Wall receptacle (See Note 1.)	Inline receptacle (See Note 1.
9811	9890	Line cord for United States, Canada, Latin America, and Japan with plug; single-phase, 60 A, two wires + ground	Hubbell HBL360P6V04	Hubbell HBL360R6W	Hubbell HBL360C6W
9813	9890	Line cord for EMEA and Asia/Pacific (except Japan) hard wired; single-phase, 60 A, two wires + ground	Hardwired or uses client-provided connector	Not applicable	Not applicable
9820	9891	Line cord for United States, Canada, Latin America, and Japan with plug; three-phase, 60 A, three wires + ground	Hubbell HBL460P9V05	Hubbell HBL460R9W	Hubbell HBL460C9W
9822	9892	Line cord for EMEA and Asia/Pacific (except Japan) hard wired; three-phase, 30 A, three wires + neutral + ground	Hardwired or uses client-provided connector	Not applicable	Not applicable
9824	9894	Line cord for United States, Canada, Latin America, and Japan with plug; three phase, 30A, three wires + ground	Hubbell HBL430P9V04	Hubbell HBL430R9W	Hubbell HBL430C9W

Note: Line cords are provided with the connector part numbers shown. Receptacle part numbers shown are recommended. Although equivalent receptacles can be used, it is the responsibility of the customer to verify compatibility.

Figure 3-14 Power cord features

For more information about the appropriate ATS and power cord selection, see the *IBM XIV Storage System Gen3 Models 281x-11x and 281x-21x Planning Guide*, SC27-5412.

UPS module complex

The uninterruptible power supplies (UPS) module complex consists of three UPS units. The UPS units work together to provide a “cache battery” function to sustain the XIV Storage System in the event of power fluctuations or loss. The UPS complex sustains the system for up to 30 seconds in the event of a fluctuation, or temporary loss of external power. If there is an extended external power failure or outage that is more than 30 seconds, the UPS module complex maintains battery power long enough to allow a safe and orderly shutdown of the XIV Storage System. The complex can sustain the failure of one UPS unit and still provide adequate protection against external power outages.

Internal UPS: Although the system is protected by a UPS for internal use, you can reduce the risk of a power outage if you connect the system to an external UPS, a backup generator, or both.

The UPS complex is designed and dedicated to supporting the components inside the XIV Storage System rack. No external equipment can be plugged into the UPS units. The UPS complex is not substitute for traditional power conditioning or sustaining equipment typically found in a data center environment and is intended to be used with existing equipment.

Important: The three UPSes in the XIV Storage System appear to the central UPS in the main data center as standard storage system power supplies. The XIV Storage System does not affect the central UPS unit in the main data center in a harmful way.

Figure 3-15 shows an illustration of one UPS module.

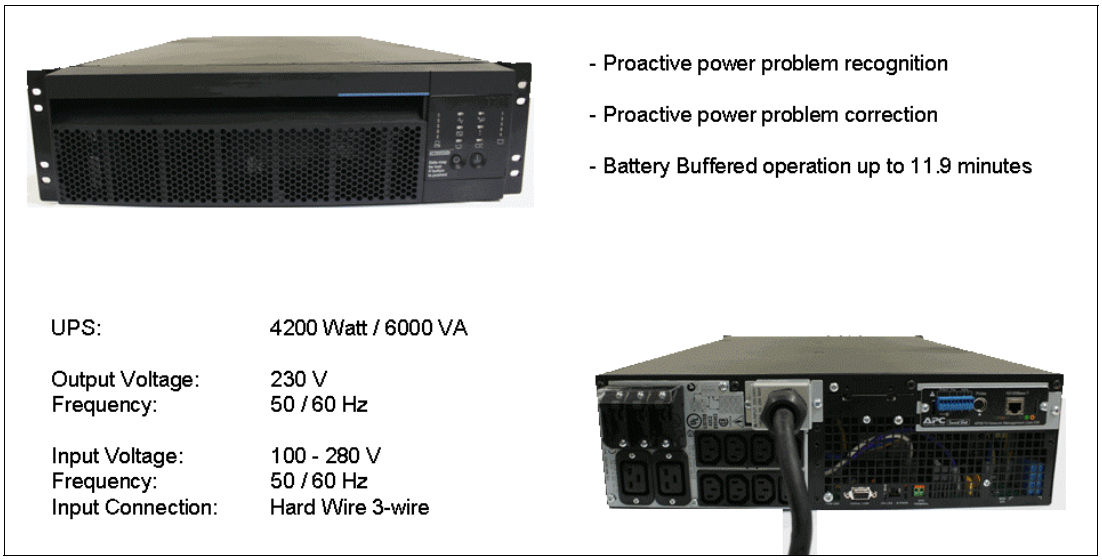

- Proactive power problem recognition
- Proactive power problem correction
- Battery Buffered operation up to 11.9 minutes
UPS: 4200 Watt / 6000 VA
Output Voltage: 230 V
Frequency: 50 / 60 Hz
Input Voltage: 100 - 280 V
Frequency: 50 / 60 Hz
Input Connection: Hard Wire 3-wire

Figure 3-15 UPS module

The three UPS modules are at the bottom of the rack. Each UPS is 3U in height. Each UPS has an output of 6 Kilovolt ampere (kVA) to supply power to all other components in the system.

Attention: Do *not* power off the XIV Storage System by using the UPS power button, because this action can result in the loss of data and system configuration information. Powering the system off must be done solely from either the XIV Storage Management graphical user interface (GUI) or the XIV command line interface (XCLI).

To monitor battery life, the UPS modules routinely run a self-test every 14 days, with a 9-hour interval between each UPS. To maximize battery life, a UPS battery calibration is performed every 120 days. This calibration drains the batteries to about 20% and recharges them to 100%. This routine operation causes UPS warning lights and audible alarms. The best way to determine whether the UPS lights and alarms are a real problem is to check the event log.

The routine calibration tests produce an event that looks similar to the event shown in Figure 3-16.

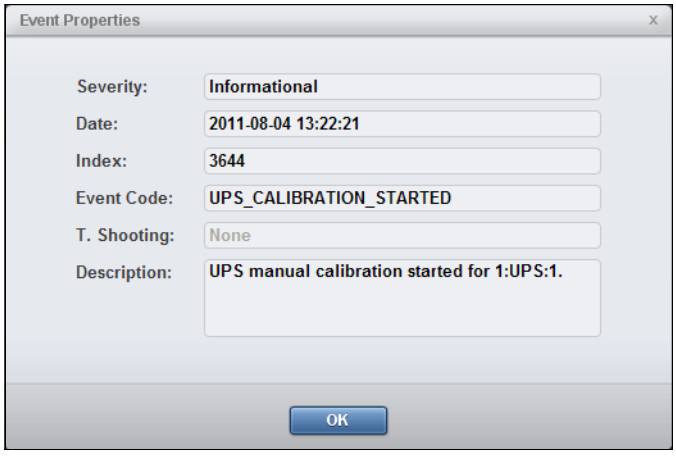


Figure 3-16 UPS calibration

3.1.9 Data modules and interface modules

The hardware of the interface modules and data modules in the XIV Gen3 is based on an Intel server platform that is optimized for data storage services. A module is 87.9 mm (3.46 inches) (2U) high, 483 mm (19 inches) wide, and 707 mm (27.8 inches) deep.

Figure 3-17 shows a representation of the inside of a data module.

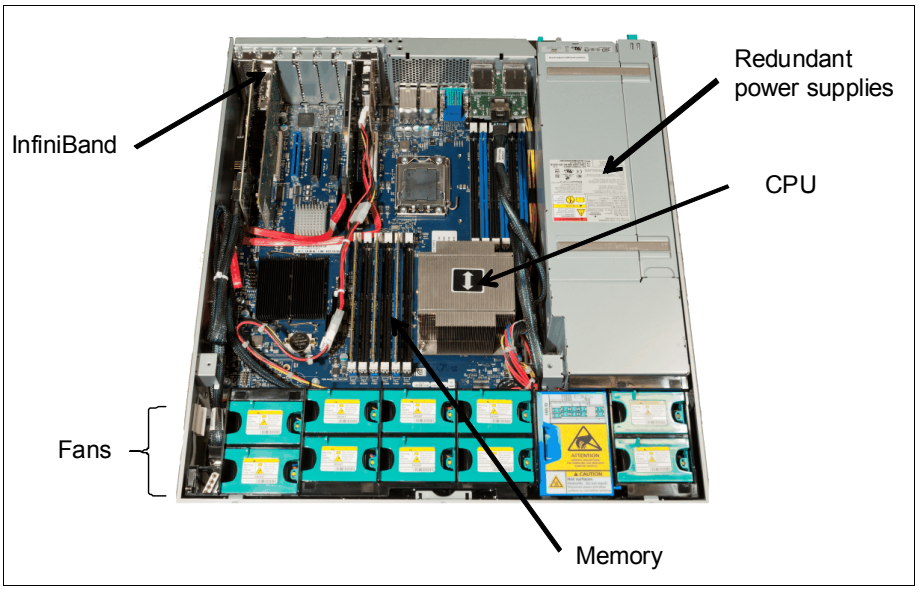


Figure 3-17 Data module

Both the data and interface modules on XIV Gen3 contain the following hardware features:

- System board with a PCIe 2.0 bus, with the following management ports (Figure 3-18 on page 83):
  - RS-232 serial port
  - USB port (one used, three unused)
  - Two Gb Ethernet ports
- Quad-core or six-core processor
- Memory
- SAS host bus adapter (HBA)
- SAS disk drives
- InfiniBand host channel adapter
- SSD slot (see Figure 3-18 on page 83)
- Fan bank
- Memory flash card (see Figure 3-18 on page 83)
- Redundant power supplies
- Enclosure management card

Figure 3-18 shows the rear of the data module with its ports and connectors.

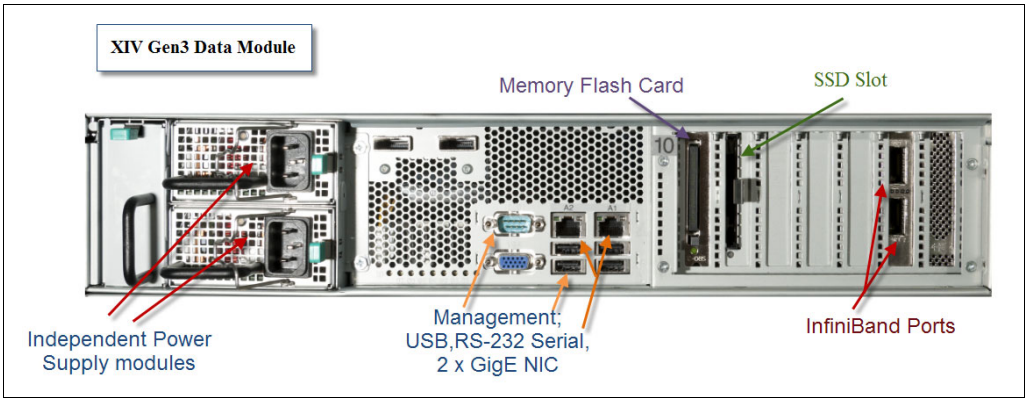


Figure 3-18 XIV Storage System Gen3 data module connections that are used

Model 114 unique components

Data modules and interface modules in the Model 114 have the components described in the subsections that follow.

Quad-core processor

The XIV Storage System Gen3 Model 114 processor is a third-generation Intel Nehalem micro-architecture Westmere 32 nm processor with a 2.4 GHz clock speed. There is one processor per module. This processor features turboboost and hyperthreading and is more energy efficient than its predecessors. This processor is specifically sized to handle the I/O workload required in the XIV Storage System Gen3 modules.

DDR3 memory model 114

XIV Storage System Gen3 has DDR3 memory that is twice the speed of its DDR2 predecessor. Each XIV Storage System Gen3 has 24 GB, which is 8 GB more memory than the model A14. DDR3 also uses less power than DDR2. At 24 GB per module, a full XIV Storage System Gen3 has 360 GB of memory.

Part of the memory is used as module system memory; the rest is used as cache memory. Cache consists of both read cache, where the module holds previously read data and pre-fetched data; plus write cache, where the module holds data that is being cached before being de-staged to disk.

When the optional SSD cache extension is installed (Model 114 software 11.1.0 or later), each module gets an extra 400 GB capacity for use as an extension of the DDR3 cache.

Power supplies

The modules are powered by a redundant power supply unit (PSU) cage with dual 850 W PSU assemblies, as shown in Figure 3-19 on page 84. These power supplies can be replaced individually with no need to stop using the module. The power supply is a field-replaceable unit (FRU). Each power supply is cabled to a different UPS unit.

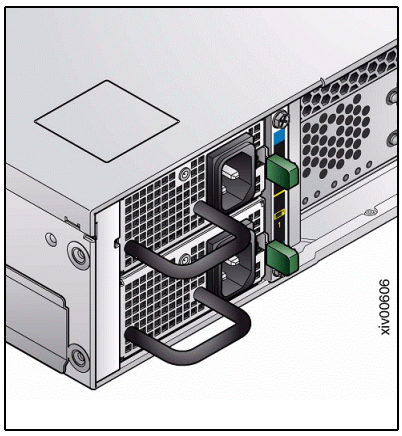


Figure 3-19 Redundant module PSUs

Model 214 unique components

Data modules and interface modules in Model 214 have the components described in the subsections that follow.

Six-core processor

Model 214 modules are powered by a six-core Intel Westmere CPU E5645 processor, which is based on the 32 nm technology and 2.4 GHz clock frequency. Equipped with 12 MByte L3 cache and DDR3 1333 MHz, the CPU features Hyper-Threading technology (HT technology), which delivers two processing threads per physical core. Consequently, a fully configured system provides the processing power of ninety physical, or 180 logical cores, using the Hyper-Treading technology. Thanks to the Integrated Power Gates, the power consumption of inactive kernels is near zero, which makes this processor highly efficient. With six cores and the mentioned feature, the processor is powerful enough to manage the basic workload of all modules and the additional workload in the interface modules.

Note: Four-core and six-core modules can be intermixed in Model 114 with software 11.2 or later, in case of FRU replacement or miscellaneous equipment specification (MES) installation.

DDR3 memory model 214

The XIV Gen3 Model 214 is equipped with up to 48 GB (3 x 16 GB dual inline memory module (DIMM)) fast DDR3 synchronous dynamic random access memory (SDRAM) per module. All modules have 48 GB installed. A 48 GB RAM is compatible with an earlier version to 24 GB RAM, so it accepts a 48 GB RAM FRU to a 24 GB RAM system but does not accept a 24 GB RAM FRU to a 48 GB RAM system. The SDRAMs are provided from two different vendors and cannot be mixed.

Power supplies

The modules in Model 214 are powered by two redundant high efficiency power supply units (HE PSU), which have reduced power consumption compared to previous standard PSUs. They are monitored by the software and can be individually replaced with no need to stop using the module. The power supply is a field-replaceable unit (FRU). Each power supply is cabled to a different UPS unit.

SAS self-encrypting disk drives

The XIV Storage System Gen3 modules can be configured with 2 TB, 3 TB, or 4 TB SAS self-encrypting drives (SEDs). These drives use the SAS 2.0 protocol and run at 7200 rpm. They incorporate industry-standard Self-Monitoring Analysis and Reporting Technology (SMART) for enhanced reliability through prediction and early recognition of drive failure conditions. The difference from normal SAS disk drives is that SAS SED drives enable the capability for encryption for data at rest. See the IBM Redpaper titled *IBM XIV Security with Data-at-Rest Encryption*, REDP-5047.

Model 114 and Model 214 common components

Model 114 and Model 214 share the components described in the sections that follow.

System board

The XIV Storage System system board uses the PCIe 2.0 bus, which is twice the speed of the XIV Storage System Gen 2 PCIe 1.0 bus and features more PCIe lanes than its predecessor.

The system board has one RS-232 serial port, one USB port, and two Ethernet ports, which are used for internal XIV Storage System management. The monitor port and three USB ports are unused.

SAS host bus adapter

The SAS host bus adapter (SAS HBA) is the controller card for the SAS disk drives in each module. This HBA uses eight lanes of the PCIe 2.0 bandwidth to communicate with the processor and cache complex. To communicate with the SAS drives, this HBA uses the SAS 2.0 standard, which is a performance improvement over its SAS 1.0 predecessor. This SAS HBA also realizes processor efficiencies because it does not have to complete a SAS to SATA conversion. SAS also outperforms SATA because unlike SATA, it features full duplex communication with each disk.

SAS disk drives

XIV Gen3 modules can be configured with 2 TB, 3 TB, 4 TB , or 6 TB SAS drives. Drive sizes cannot be mixed within the same XIV Storage System. These drives use the SAS 2.0 protocol and run at 7200 rpm. They incorporate SMART for enhanced reliability through prediction and early recognition of drive failure conditions. The drive has data prefetch and cache (drive buffer storage) control algorithms that improve drive performance. These drives are also capable of the following reliability features:

- Background media scan to detect drive media errors
- Media pre-scan to allow the drive to repair media errors
- Deferred auto-reallocation, which allows the drive to reallocate unreadable locations on a subsequent write command
- Idle read after write, which uses drive idle time to verify the integrity of recently written data

This SAS drive also exhibits significant performance gains for large block I/O.

All XIV Storage System disks are installed in the front of the modules, with 12 disks per module. Each single SAS disk is installed in a disk tray that connects the disk to the backplane and includes the disk indicators on the front. If a disk is failing, it can be replaced easily from the front of the rack. The complete disk tray is one FRU, which is latched in its position by a mechanical handle.

Important: IBM performs all disk replacements in an XIV Storage System. SAS disks in the XIV must never be swapped within a module or placed in another module because of the internal tracing and logging data that they maintain.

In the XIV Gen3, disks are replaced according to the IBM standard maintenance strategy in “deferred maintenance”. This means that IBM is alerted with the third failing disk in the system and arranges the replacement of all three failing disks in one repair action by the IBM SSR. Optionally, with extended maintenance, each failing disk is replaced immediately after it was failed.

InfiniBand host channel adapter

The InfiniBand host channel adapter (HCA) is a 2x double data rate (DDR) host channel adapter that features 20 Gbps of internal bandwidth and has low latency. The two ports are cabled to redundant InfiniBand switches inside the XIV Storage System Gen3 rack. The InfiniBand HCA has the task of protocol processing, which offloads this workload from the module’s processor and communicates with the processor complex over eight lanes of PCIe 2.0 bus bandwidth. InfiniBand does extensive end-to-end link level protocol checking, so it is reliable.

Flash cache

One of the most important features of the XIV Gen3 is the upgrade with flash cache. The optional flash cache option is available for Gen3 Model 214 and Model 114 running software Version 11.1.0 or later. XIV Gen3 systems can be equipped with a 400 GB SSD for every module, and the XIV Gen3 with 4 TB or 6 TB drives can be equipped with a 400 GB or 800 GB SSD for every module. The installation is nondisruptive. For a 15-module system, this configuration represents 6 TB of read flash cache with the 400 GB SSD feature, and with 800 GB SSD, it has 12 TB read flash cache. For partial configurations, see Figure 3-2 on page 65.

There is not an additional disk tier to manage; the XIV Storage System Gen3 software manages this flash cache automatically.

This flash cache feature further improves performance for read workloads, especially random read workloads.

Memory flash card

Each module contains a CompactFlash card, which is shown in Figure 3-20. Also see Figure 3-18 on page 83 and Figure 3-22 on page 88.

Figure 3-20 Module CompactFlash card

This card is the boot device of the module and contains the software and module configuration files.

Important: Because of the configuration files, the Compact Flash Card is not interchangeable between modules, and it is not user-serviceable.

Cooling fans

To provide enough cooling for the disks, processor, and system board, the system includes 10 fans located between the disk drives and the system board. Cool air is pulled from the front of the module through the disk drives. The air flow and the alignment of the fans ensures the appropriate cooling of the entire module, even if a fan is failing.

Enclosure management card

The enclosure management card is located between the disk drives and the system board. In addition to the internal module connectivity between the drive backplane and the system board, this card is the backplane for the 10 fans. It also includes the fan control and the logic to generate hardware alarms in the case of problems in the module.

3.1.10 Interface modules

From a hardware perspective, the XIV Gen3 interface modules are the same as the data modules in the corresponding Model 114 or Model 214. See the diagram in Figure 3-21 on page 87. Additional hardware, Fibre Channel HBAs, and Ethernet adapters differentiate the interface module from a data module.

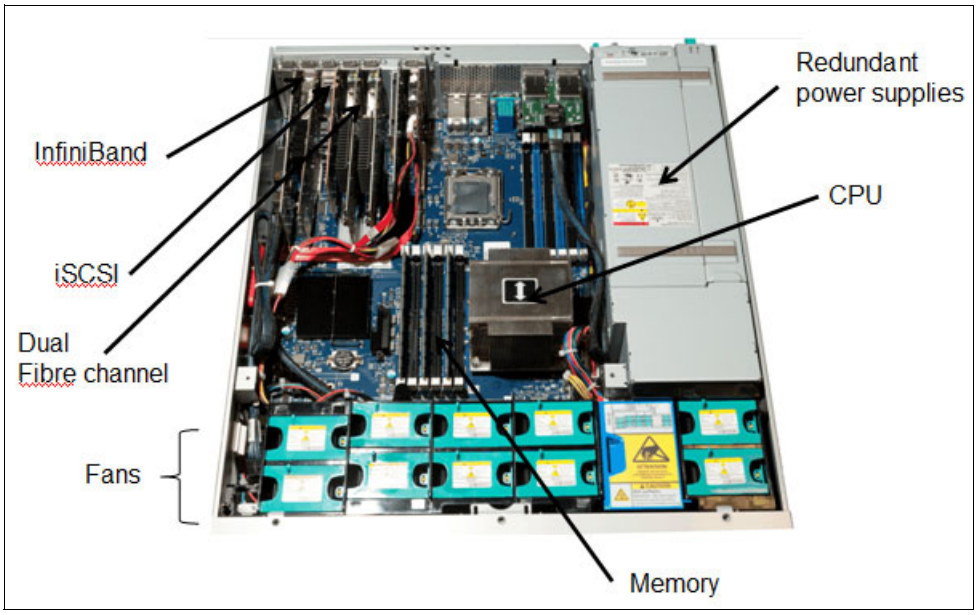


Figure 3-21 Interface module

Depending on the Gen3 model, a different number of adapters are installed in the interface modules:

- Model 114 and 214 (1 GbE):
  - Two 2-port 8 Gbps Fibre Channel HBAs
  - One 4-port iSCSI 1 Gbps Ethernet adapter
- Model 214 (10 GbE):
  - Two 2-port 8 Gbps Fibre Channel HBAs
  - One 2-port iSCSI 10 Gbps Ethernet adapter

These two interface adapter types are used for host attachment to the XIV Storage System. The ports can also be used to establish remote mirror links and data migration paths with another remote XIV or storage systems from other vendors (migration).

For the number of usable iSCSI ports in XIV Gen3 partial configurations, see Figure 3-2 on page 65.

Figure 3-22 on page 88 shows a schematic (rear view) of the two interface modules with Fibre Channel and iSCSI ports. All Fibre Channel and iSCSI ports used for external connections are internally connected to a patch panel, where the external connections are made. The patch panel layout depends on the XIV model. See 3.1.12, “Patch panel” on page 92, which shows the two different patch panels.

Figure 3-22 shows an XIV Model 114/214 (1 GbE) and XIV Model 214 (10 GbE) interface module with Fibre Channel and iSCSI ports.

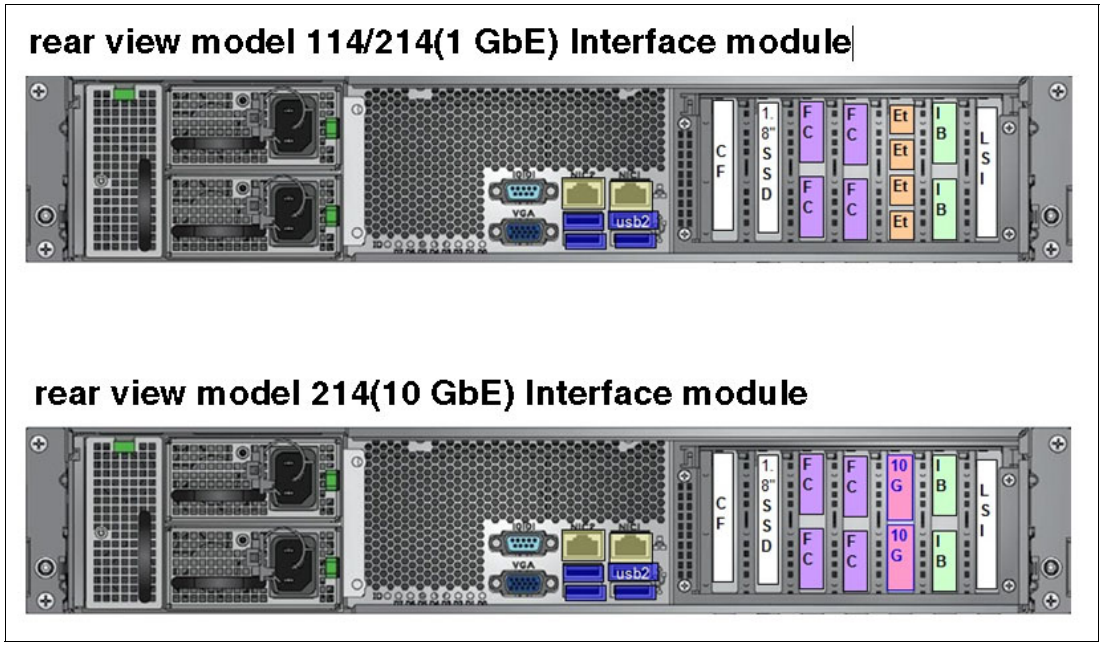


Figure 3-22 XIV Model 114/214 (1 GbE) and Model 214 (10 GbE) interface module with FC/iSCSI ports

Fibre Channel connectivity

Each XIV Gen3 interface module has two 2-port 8 Gbps HBAs from Emulex. Each HBA has eight PCIe 2.0 lanes to the processor complex. The HBAs feature hot-upgradable firmware capability, and they support N-Port ID Virtualization (NPIV) and 1400 concurrent exchanges per port.

There are four Fibre Channel ports available in each interface module, for a total of 24 Fibre Channel ports in a fully configured XIV system. Certain partial rack configurations do not use all ports, even though they might be physically present.

The Fibre Channel ports support 2, 4, and 8 Gbps full-duplex data transfer over short wave fibre links, using 50 micron multi-mode cable. It is not possible to attach this HBA to a 1 Gbps SAN switch.

The allocation and use of specific ports in each module depend on your environment, your specific requirements in terms of resiliency, the nature of your host I/O traffic, and whether you use mirroring or not.

Most illustrations in this book show ports 1 and 3 allocated for host connectivity. Ports 2 and 4 can be reserved for additional host connectivity or remote mirror and data migration connectivity. This configuration is generally the choice for clients who want more resiliency

(ports 1 and 3 are on separate adapters), or availability (in case of adapter firmware upgrade, one connection remains available through the other adapter). From a performance perspective, each adapter has its own performance bandwidth capabilities and dedicated PCIe bus bandwidth. If you do not use mirroring, you can also change port 4 to a target port.

Fibre Channel ports: Using more than 12 Fibre Channel ports for host connectivity does not necessarily provide more bandwidth. A preferred practice is to use enough ports to support multipathing, without overburdening the host with too many paths to manage. See *IBM XIV Storage System: Host Attachment and Interoperability*, SG24-7904 for more information.

iSCSI connectivity

With the 10 GbE adapter that equips the Model 214 (10 GbE), there are differences in iSCSI connectivity between the Model 114, Model 214 (1 GbE), and Model 214 (10 GbE).

In addition, improvements in the XIV Storage Software Version 11.2 include a new iSCSI implementation with significant performance enhancements:

- Network protocol processing has been moved from the Linux kernel to a user space application based on the open source Lightweight IP (LWIP) stack. This change results in significant CPU processing efficiencies.
- Using multiple Transmission Control Protocol (TPC) connections (iSCSI sessions) to provide higher throughput.
- Provides more robust host connectivity during XIV hot upgrade.

Model 114 and Model 214 (1 GbE)

There are up to 22 iSCSI ports available for iSCSI over IP/Ethernet services. The exact port count depends on the number of active interface modules. See Figure 3-2 on page 65 for details about active interface modules. These ports support 1 Gbps Ethernet network connection each. These ports connect to the user's IP network through the XIV patch panel and provide connectivity to the iSCSI hosts. See Figure 3-26 on page 92 for more details about the cabling of these ports.

You can operate iSCSI connections for various uses:

- As an iSCSI target that attached hosts can connect to using the iSCSI protocol
- As an iSCSI initiator for remote mirroring when connected to another iSCSI port
- As an iSCSI initiator for data migration when connected to a iSCSI storage system from another vendor

For each iSCSI IP interface, you can define these configuration options:

- IP address (mandatory)
- Network mask (mandatory)
- Default gateway (optional)

The default and highest possible maximum transmission unit (MTU) is 9000 MTU.

Model 214 (10 GbE)

In Model 214 (10 GbE), slot 3 of the interface module is equipped with a 10 GbE adapter. The adapter is a Mellanox ConnectX-2 EN network interface card, as shown in Figure 3-23 on page 90. It serves two 10 GbE ports using Optical Transceiver Modules (SFP+).

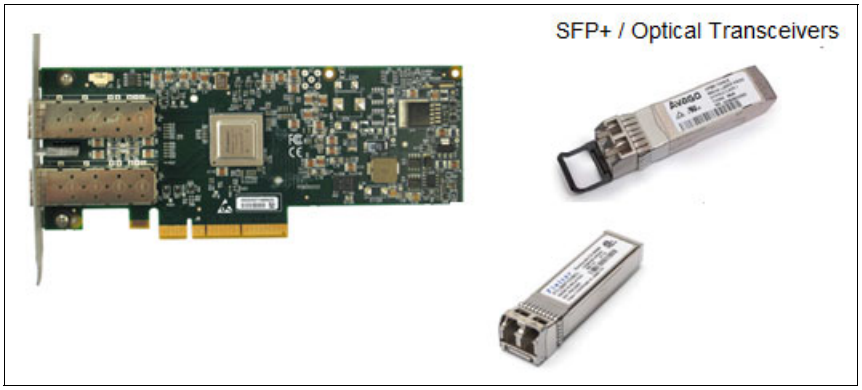


Figure 3-23 Model 214 10 GbE adapter

The 10 GbE adapter provides enhanced iSCSI host connectivity and adequate infrastructure for potential FCoE (Fibre Channel over Ethernet) functions, if IBM decides to offer FCoE.

There are up to twelve 10 GbE ports available for iSCSI over IP/Ethernet services. The number of active ports depends on the number of modules installed in the rack. See Figure 3-2 on page 65. The active ports are connected to the client's IP network through the patch panel. For that reason, the patch panel in Model 214 (10 GbE) differs from the patch panel in the previous Model 114 and Model 214 (1 GbE), as shown in Figure 3-26 on page 92.

You can operate iSCSI connections for various uses:

- As an iSCSI target that attached hosts can connect to using the iSCSI protocol
- As an iSCSI initiator for remote mirroring when connected to another iSCSI port
- As an iSCSI initiator for data migration when connected to an iSCSI storage system from another vendor

For each iSCSI IP interface, you can define these configuration options:

- IP address (mandatory)
- Network mask (mandatory)
- Default gateway (optional)

The default and highest possible MTU is 9000 MTU.

Note: With Model 214 (10 GbE), the number of iSCSI ports is divided in half, per interface module. However, because there are 10 GbE ports, the iSCSI bandwidth is five times greater than with Model 114 and Model 214 (1 GbE).

3.1.11 InfiniBand module interconnect

The internal network is based on two redundant 36-port InfiniBand switches. InfiniBand is a switch fabric architecture characterized as fast with low latency.

Each of the modules (data or interface) has an InfiniBand HCA that is cabled to each of the InfiniBand switches. The switches are also linked to each other. See Figure 3-24 for a logical view of this connectivity. For external views of the adapter and its ports, see Figure 3-18 on page 83, Figure 3-21 on page 87, and Figure 3-22 on page 88.

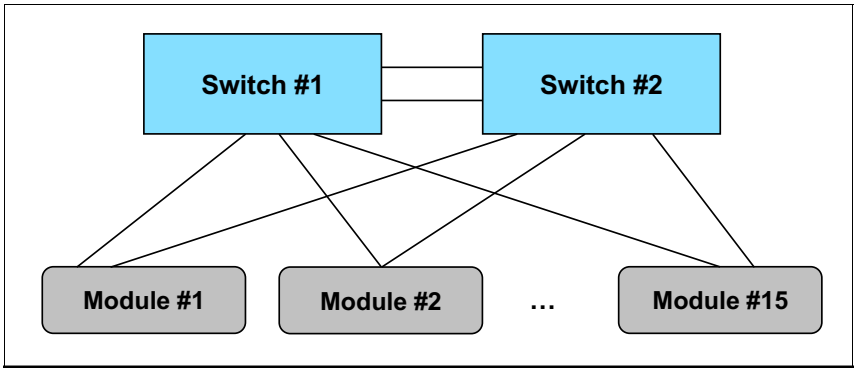


Figure 3-24 InfiniBand internal cabling

This network topology enables maximum bandwidth use because the switches are used in an active-active configuration. The InfiniBand switches are also tolerant to any failure of the following individual network components:

- Ports
- Links
- Switches

Figure 3-25 shows the two InfiniBand switches installed in the XIV rack. The cabling has been removed from the top InfiniBand switch to show a better view of the switch itself.

Figure 3-25 Model 114 dual 36-port InfiniBand switches

Each InfiniBand switch contains 36 ports that have 40 Gbps full bidirectional bandwidth per port. Port-to-port latency is less than 100 nanoseconds. Each switch has 2.88 Tbps switching throughput. The switches are powered by redundant power supplies and fan modules to eliminate any single point of failure. Additionally, each switch has several RJ-45 management ports that are used by the XIV Storage System. InfiniBand is also scalable, well beyond the current 15 modules in a single XIV Storage System rack.

3.1.12 Patch panel

The patch panel is at the rear of the rack. The Fibre Channel ports on the interface modules are connected to the patch panel using 50-micron cables. All external connections must be made through the patch panel. In addition to the host connections and to the network connections, more ports are available on the patch panel for service connections. Figure 3-26 shows the details for the patch panel and the ports.

The patch panel has had various redesigns of the labeling based on the production date.

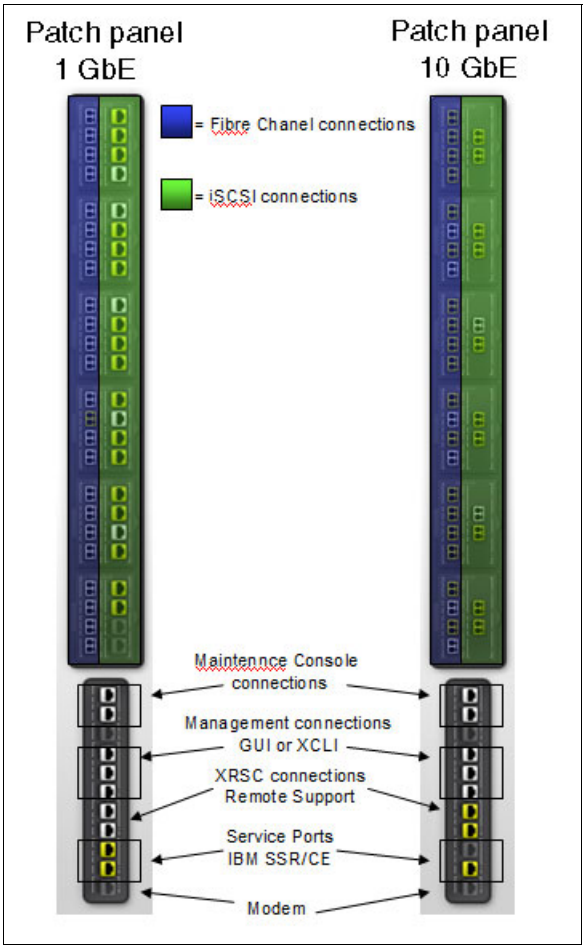


Figure 3-26 Connection details of patch panel ports

3.1.13 Hardware support

This section covers important features of the XIV Storage System that are used by internal functions and IBM maintenance technicians if a problem arises with the system.

Module USB-to-serial connections

The module USB-to-serial connections are used by internal system processes to keep the inter-communication between the modules active when the normal network connection is not operational. Modules are linked together with USB-to-serial cables in groups of three. These links are also used by IBM Hardware Maintenance Services for repair to the internal network. The connection sequence is shown in Figure 3-27.

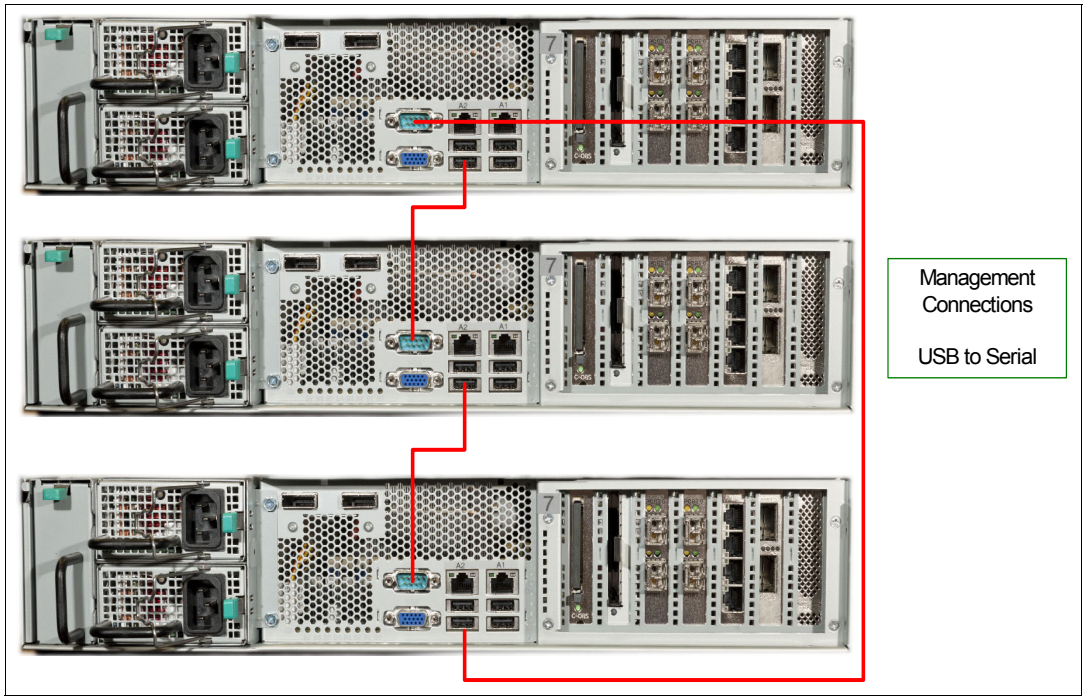


Figure 3-27 Module: USB to serial

Maintenance module

The 1U maintenance module and the modem, which are installed in the middle of the rack, are used for XIV Storage System support and for the IBM personnel to maintain and repair the system. This device is used only to gain remote access to the XIV System through the modem for support personnel. When there is a software or hardware problem that needs attention, a remote connection is required and used to analyze and possibly repair the faulty system. If there is no XIV Remote Support Center (XRSC) or other broadband connection available, the only way to connect is through the modem and maintenance module. For more information about remote connections, see “XIV Remote Support Center” on page 101.

Modem

The modem installed in the rack is an option to use for remote support if the preferred choice of XIV Remote Support Center is not selected. It enables the IBM specialists and higher-level support to connect to the XIV. Problem analysis and repair actions without a remote connection can be complicated and time-consuming.

Modem: The modem is not available in all countries.

3.2 Hardware planning overview

This section provides an overview of planning considerations for the XIV Storage System, including a reference listing of the information that is required for the setup. The following information is included in this section:

- Physical installation requirements
- Delivery requirements
- Site requirements
- Cabling requirements

For more detailed planning information, see the *IBM XIV Storage System Gen3 Models 281x-11x and 281x-21x Planning Guide*, SC27-5412.

Additional documentation is available in IBM XIV Storage System documentation in the IBM Knowledge Center:

http://ibm.co/1rvfciG

For a smooth and efficient installation of the XIV Storage System, planning and preparation tasks must take place before the system is scheduled for delivery and installation in the data center. A sales representative will arrange a Technical Delivery Assessment (TDA) meeting to go over site-specific details and to ensure that the correct information is gathered before the delivery of the system.

Installation planning involves the following major areas:

- Ordering the appropriate hardware configuration:
  - Current requirements
  - Future requirements
- Configuration planning:
  - Basic configurations
  - Network connections
  - Management connections
- Physical installation

3.2.1 Basic configuration planning

You must complete the configuration planning worksheets and checklists from the Planning topic of the IBM XIV Storage System documentation in the IBM Knowledge Center to allow the IBM Service Support Representative (SSR) to physically install and configure the system (IBM XIV Storage System > IBM XIV Gen3 281x-11x and 281x-21x > Planning):

http://ibm.co/1raoi5t

The configuration planning worksheets are also in the *IBM XIV Storage System Gen3 Models 281x-11x and 281x-21x Planning Guide*, SC27-5412.

Additional documentation is available from the XIV Storage System Information Center at:

http://publib.boulder.ibm.com/infocenter/ibmxiv/r2/index.jsp

You must provide the IBM SSR with the information required to attach the system to your network for operations and management, and enable remote connectivity for IBM support and maintenance.

Enter all of the following required information in each worksheet to prevent further inquiry and delays during the installation (see 3.2.4, “IBM XIV Storage System physical installation” on page 105):

- Interface module:

  Interface modules (4, 5, and 6) need an IP address, netmask, and gateway. This address is necessary to manage and monitor the XIV Storage System by using either the XIV Storage Management GUI or the XCLI. Each interface module needs a separate IP address in case a module is failing.

- Domain name server:

  If Domain Name System (DNS) is used in your environment, the XIV Storage System must have the IP address, netmask, and gateway from the primary DNS server and, if available, the secondary server.

- Simple Mail Transfer Protocol gateway:

  The Simple Mail Transfer Protocol (SMTP) gateway is needed for event notification through email. The XIV Storage System can initiate an email notification, which is sent out through the configured SMTP gateway (IP address or server name, netmask, and gateway).

- Network Time Protocol (time server):

  XIV Storage System can be used with a Network Time Protocol (NTP) time server to synchronize the system time with other systems. To use this time server, the IP address (or server name), netmask, and gateway must be configured. It is a preferred practice to enter the information for an NTP server.

- Time zone:

  Usually the time zone depends on the location where the system is installed. However, exceptions can occur for remote locations where the time zone equals the time of the host system location.

- Email sender address:

  The email address to show as the sender in the email notification.

- Remote access:

  The modem number or a client-side IP address must be configured for remote support. This network connection must have outbound connectivity to the Internet.

This basic configuration data is entered in the system by the IBM SSR after the physical installation. See “Basic configuration” on page 105.

Other configuration tasks, such as defining storage pools, volumes, and hosts, are the responsibility of the storage administrator and are described in Chapter 4, “IBM XIV Storage Management software” on page 111.

Network connection considerations

Network connection planning is also essential to installing XIV Storage System. To deploy and operate the system in your environment, several network connections are required:

- Fibre Channel connections for host I/O over Fibre Channel
- 1 Gigabit Ethernet connections for host I/O over iSCSI
- 10 Gigabit Ethernet connections for host I/O over iSCSI
- Gigabit Ethernet connections for management
- Gigabit Ethernet connections for IBM XIV Storage System remote support

All external XIV Storage System connections are through the patch panel, as explained in 3.1.12, “Patch panel” on page 92.

For details about the host connections, see *XIV Storage System: Host Attachment and Interoperability*, SG24-7904.

Fibre Channel connections

When shipped, the XIV Storage System is equipped by default with 24 Fibre Channel ports (assuming a fully populated 15 module rack). The XIV Storage System supports 50-micron fiber cabling. If you have other requirements or special considerations, contact your IBM SSR.

The 24 Fibre Channel (FC) ports are available from the six interface modules, four in each module, and they are internally connected to the patch panel. Of the 24 ports, 12 are provided for connectivity to the switch network for host access, and the remaining 12 are for use in remote mirroring or data migration scenarios (however, they can be reconfigured for host connectivity). Adhere to this guidance on Fibre Channel connectivity. The external (client-provided) cables are plugged into the patch panel. For planning purposes, Figure 3-28 highlights the maximum values for various Fibre Channel parameters for your consideration.

FC parameters	Maximum values
	Modell 114, Model 214(1 GbE) and Model 214(10 GbE)
Maximum number of Interface Modules with iSCSI ports	6
Maximum number of 8-GB FC ports	24
Maximum queue depth per FC host port	1400
Maximum queue depth per mapped volume per (host port, target port, volume)	256
Maximum FC ports for host connections (default configuration)	12
Maximum FC ports for mirror/migration connections (default configuration)	12
Maximum number of concurrent host logins / FC adapter	350
Maximum number of concurrent host logins / Interface Module	700
Maximum number of concurrent host logins / 15 Module XIV	2000
Maximum number of hosts ports defined (WWPN + IQNs)	4000

Figure 3-28 Fibre Channel parameters

Values in Figure 3-28 apply to Version 11.5 of the XIV Storage System Software.

iSCSI connections

The external (client-provided) Ethernet cables (Model 114 and Model 214 [1 GbE]) or optical cables (model 214 [10 GbE]) are plugged into the patch panel. For planning purposes, highlight the maximum values for various iSCSI parameters for your consideration. As with Fibre Channel, it is important to plan your connectivity based on these maximums.

iSCSI parameters	Maximum values		
	Modell 114	Model 214(1 GbE)	Model 214(10 GbE)
Maximum number of Interface Modules with iSCSI ports	6	6	6
Maximum number of 1-GB iSCSI ports(Interface Module 4 only have 2 ports)	22	22	-
Maximum number of 10-GB iSCSI ports	-	-	12
Maximum number of hosts ports defined (WWPN + IQNs)	4000	4000	4000

Figure 3-29 iSCSI parameters

These values are correct at the time of publication of the Ninth Edition of this book for version 11.5 of the XIV Storage System Software.

Host multipathing

There are two main multipathing techniques. For both multipathing configurations, each host engages the I/O services of every XIV Storage System interface module.

Experience has shown that the first multipathing configuration in Figure 3-30 is the best overall general-purpose configuration. Host multipathing reliability during path error recovery in some operating systems is complicated by increasing numbers of paths per logical unit number (LUN). Certainly, for host systems with two HBAs, using the six paths per LUN is a preferred practice.

Tip: Six paths per LUN is the preferred multipathing configuration.

There is no performance or reliability benefit in using too many paths. Going beyond 12 paths per volume has no benefit, because most operating systems have more processing overhead handling the paths and cause longer times for recovery. Going beyond six paths rarely has much benefit. Use four or six paths per volume as a standard. Do not go below four paths.

Figure 3-30 shows six paths per LUN as the best multipathing configuration.

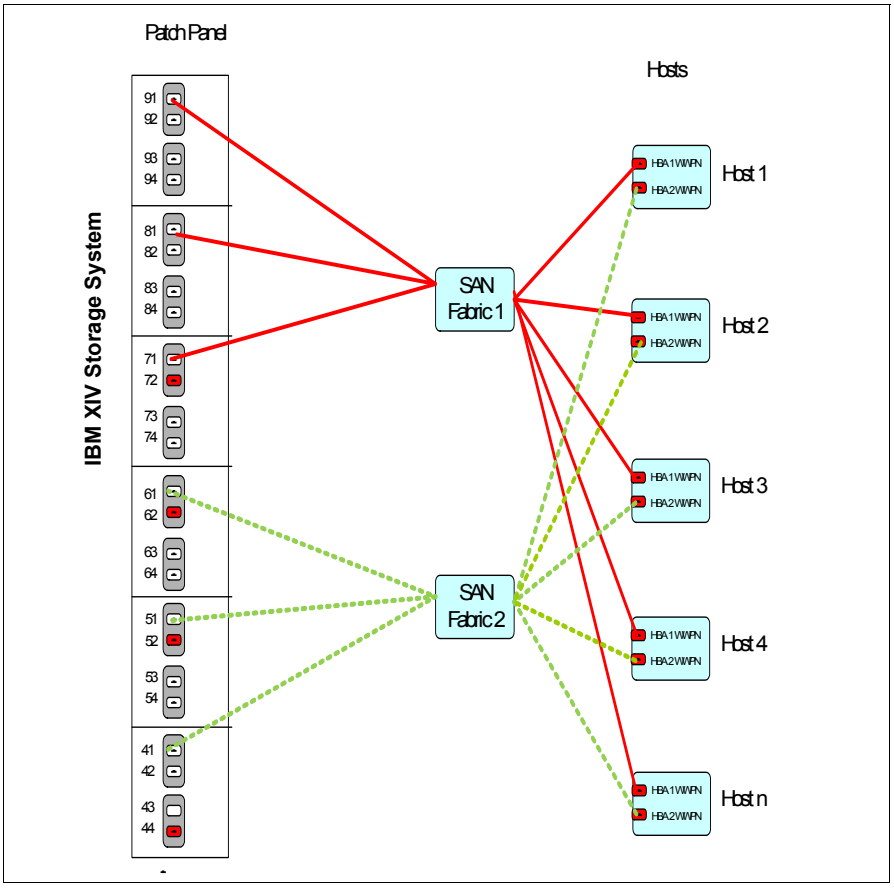


Figure 3-30 Six paths per LUN is the best overall multipathing configuration

The following information relates to Figure 3-31 on page 98:

- Each host is equipped with dual HBAs. Each HBA (or HBA port) is connected to one of two FC switches.
- Each of the FC switches has a connection to three separate interface modules.

- Each LUN has six paths.

The second multipathing configuration that is shown in Figure 3-31 on page 98 is more appropriate for benchmarking and higher performance host systems with the highest I/O requirements. The primary differentiator is the host's ability to handle the higher number of paths per LUN. Do not use the configuration in Figure 3-31 on page 98 for most production applications, because the host multipathing provides enough reliability during path recovery scenarios.

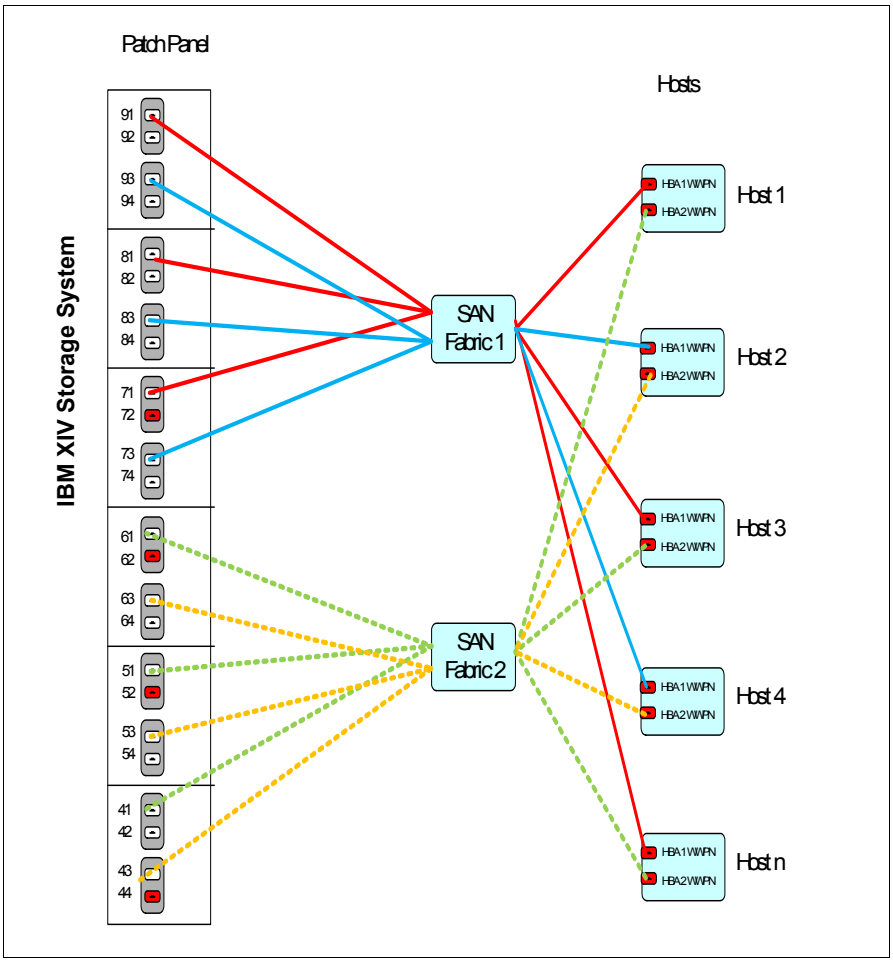


Figure 3-31 Multipathing configuration for larger hosts

The following information relates to the configuration that is shown in Figure 3-31:

- Each host is equipped with dual HBAs. Each HBA (or HBA port) is connected to one of two FC switches.
- Each of the FC switches has a connection to a separate FC port of each of the six interface modules.
- Each LUN has 12 paths.

Each host HBA is physically connected to all six of the XIV Storage System interface modules. This configuration provides the ability for each HBA to use the maximum available I/O capabilities.

For XIV Storage System, each interface module has two 2-port Fibre Channel adapters. It is a preferred practice that each zone is physically connected to one port on each XIV Storage

System Fibre Channel adapter. For example, switch A's zone is connected to port 1 on the interface modules and switch B's zone is connected to port 3 on the interface modules.

For more information, see the IBM Redbooks publication titled *IXIV Storage System: Host Attachment and Interoperability*, SG24-7904.

Single-switch solution

This configuration is resilient to the failures of a single interface module, host bus adapter, and cables, but in this configuration, the switch represents a single point of failure. If the switch goes down because of a hardware failure or because of a software update, the connected hosts lose all data access. Figure 3-32 on page 99 shows this configuration.

Use a single-switch solution only when no second switch is available or for test environments.

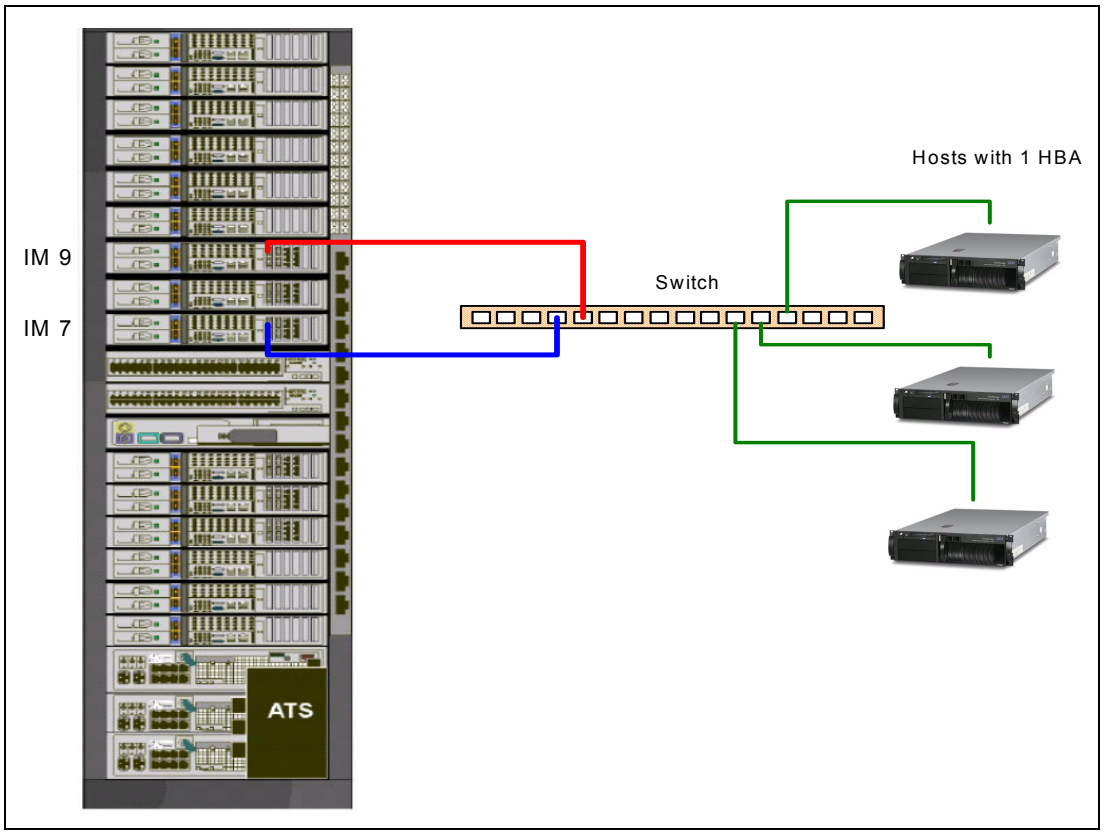


Figure 3-32 Non-redundant configuration

Single HBA host connectivity

Hosts that are equipped with a single Fibre Channel port can access only one switch. This configuration is resilient to the failure of an individual interface module, but there are various possible points of failure (switch, cable, and HBA) that can cause access loss from the host to the XIV Storage System. This configuration, which is not desirable for any production system, is used if there is no way of adding a second Fibre Channel port to the host.

Restriction: Direct host to XIV Storage System connectivity is *not* supported. The implementation must use a SAN fabric (either single or dual SAN switches). Dual fabric configurations are preferable.

Fibre Channel cabling and configuration

Fibre Channel cabling must be prepared based on the required fiber length and the selected configuration.

When installing an XIV Storage System, complete the following Fibre Channel configuration procedures:

- You must configure Fibre Channel switches that are zoned correctly to allow access between the hosts and the XIV Storage System. The specific configuration to follow depends on the specific Fibre Channel switch. It is best to have a separate zone for each initiator.
- Hosts must be set up and configured with the appropriate multipathing software to balance the load over various paths. For multipathing software and setup, see the specific operating system section in *XIV Storage System: Host Attachment and Interoperability*, SG24-7904.

iSCSI network configurations

Logical network configurations for iSCSI are equivalent to the logical configurations that are suggested for Fibre Channel networks. The following options are available:

- Redundant configuration:

  Each module connects through two ports to two Ethernet switches, and each host is connected to the two switches. This design provides a network architecture resilient to a failure of any individual network switch or module.
- Single switch configuration:

  A single switch connects all modules and hosts.
- Single port host solution:

  Each host connects to a single switch, and a switch is connected to two modules.

IP configuration

The configuration of the XIV Storage System iSCSI connection is dependent on your network. In the high availability configuration, the two client-provided Ethernet switches used for redundancy can be configured as either two IP subnets or as part of the same subnet. The XIV Storage System iSCSI configuration must match the client's network. You must provide the following configuration information for each Ethernet port:

- IP address
- Netmask
- MTU (optional, depending on your network's MTU)

  MTU configuration is required if your network supports an MTU that differs from the standard one. The largest possible MTU must be specified. From the XIV Storage System, with software Version 11.2 or later, the largest MTU size is increased to 9000. If the iSCSI hosts are on another subnet than the XIV Storage System, a default IP gateway per port must be specified.

  Important: An XIV Storage System running on software Version 11.2 or later has an increased largest-possible MTU size of 9000.

- Default gateway (optional):

  Because the XIV Storage System always acts as a Transmission Control Protocol (TCP) server for iSCSI connections, packets are always routed through the Ethernet port from

which the iSCSI connection was initiated. The default gateways are required only if the hosts are not on the same Layer-2 subnet as the XIV Storage System.

The IP network configuration must be ready to ensure connectivity between the XIV Storage System and the host before the physical system installation:

- Ethernet virtual local area networks (VLANs), if required, must be configured correctly to enable access between hosts and the XIV Storage System.
- IP routers (if present) must be configured correctly to enable access between hosts and the XIV Storage System.

Mixed iSCSI and Fibre Channel host access

The XIV Storage System does not support mixed concurrent access from the same host to the same volume through FC and iSCSI simultaneously. A single host could access different volumes where some volumes are accessed using FC and some are accessed using iSCSI. You could also access the same volume from separate hosts using various connection methods (if those hosts use a form of clustered file or operating system).

IBM service ports

The XIV Storage System has two Ethernet ports dedicated for the use of IBM service personnel. The system serves as a Dynamic Host Configuration Protocol (DHCP) server and automatically configures the notebook. These ports are only for IBM maintenance personnel, for example, an IBM Service Support Representative (SSR) or IBM Technical Advisor (TA).

Restriction: Do not connect these ports to the user (client) network.

XIV Remote Support Center

For remote support by XIV Storage System personnel, you configure a dedicated Ethernet port for remote access. This port must be connected through your organization's firewall so that IBM personnel can access the XIV, if necessary.

The XIV Remote Support Center (XRSC) merges XIV internal functions with a set of globally deployed supporting servers to provide secure IBM support access to the XIV Storage System when necessary and when authorized by the IBM client's personnel.

Figure 3-33 provides a representation of the data flow of the XIV Storage System to IBM Support.

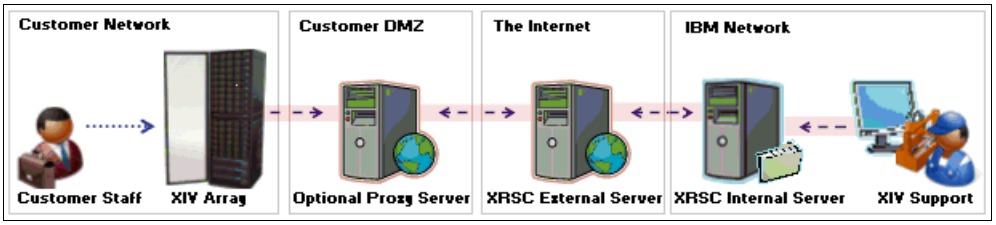


Figure 3-33 XIV Remote Support Center

An optional Remote Support Proxy can be used when one or more IBM XIV systems do not have direct access to the Internet (for example, because of firewall restrictions). You can use the XIV Remote Support Proxy to facilitate the connection to the XIV Remote Support Center. For more information about the Remote Support Proxy, see the *Installation and User's Guide*, GA32-0795.

To initiate the remote connection process, complete the following steps:

1. Initiate an Internet-based Secure Shell (SSH) connection to XRSC by using either the XIV GUI or XCLI.
2. XRSC identifies the XIV Storage System and marks it as “connected.”
3. IBM Support connects to XRSC using SSH over the IBM intranet.
4. XRSC authenticates IBM Support with the IBM intranet.
5. XRSC then shows the connected client system available to IBM Support.
6. IBM Support then chooses which system to support and connect to. Only allowed IBM XIV systems are shown. IBM Support logs the intended activity.
7. A fully recorded support session starts.
8. When complete, IBM Support terminates the session and the XRSC disconnects the XIV Storage System array from the remote support system.

The XRSC Internet servers are hardcoded in XIV Storage System Software. Therefore, no further configuration is required by the client to enable this function, aside from turning this feature on using the XIV Storage Management GUI or XCLI. This service provides an expedient means for IBM Support to gather required information from the system in the most timely fashion and with the least impact to the client.

Modem connectivity

The modem installed in the rack is optionally used for remote support if the preferred choice of XIV Remote Support Center is not used. It enables the XIV System Storage Support Center specialists and, if necessary, a higher level of support, to connect to the XIV Storage System. Problem analysis and repair actions without a remote connection can be complicated and time-consuming.

Remote mirroring connectivity

Planning the physical connections also includes considerations when the XIV Storage System is installed in a Remote Copy environment. Contact advanced IBM XIV Storage System support for assistance in planning remote mirroring connectivity to ensure maximum resilience to hardware failures and connection failures.

Remote Copy links, which connect the direct primary system and secondary system, must also be planned for before the physical installation. The physical Remote Copy links can be Fibre Channel links, direct or through a SAN, or iSCSI port connections using Ethernet. However, iSCSI is not the best option for this use.

Planning for growth

Consider growth and the future I/O demands of your business. Most applications and databases grow quickly and the need for greater storage capacity increases rapidly. Planning for growth before the implementation of the first XIV Storage System in the environment can save time and effort in the future.

3.2.2 IPv6 addressing and planning

Internet Protocol Version 6, or IPv6, is intended to supplement and, over time, replace the IPv4 protocol that is still largely used today. Getting ready to make the transition to IPv6 is becoming critical for some companies because the pool of available IPv4 addresses is already depleted. IPv6 is described by the Internet Engineering Task Force (IETE), RFC2460.

IPv6 addressing

The IBM XIV with system software Version 11.1.0 or later supports both IPv4 and IPv6 internet addressing.

IPv6 provides the following advantages over IPv4:

- Larger address space: 128 bit (2^128 ~ 3.4×10^38 addresses)
- Fixed subnet size 64 bits (2^64 (~18.4 quintillion) addresses
- Eliminates the need for network address translation (NAT)
- Supports stateless auto-configuration and network renumbering
- Inter-operability with IPv4 by fixed subnet size (64-bit addresses)
- Built in protocol for multicasting
- Mandatory support for IPSec

IPv6 compared to IPv4 addressing scheme

The larger IPv6 addressing requires a change in the way IP addresses are represented:

- IPv4 architecture:
  - Uses 32-bit addressing with four groups of decimal numbers (1 - 254) separated by dots (.), such as this example:

    `9.54.100.124`
- IPv6 architecture:
  - Uses 128-bit addressing with eight groups of four hexadecimal digits (representing 16 bits) separated by a colon (:), such as this example:

    `2001:0DB8:0000:0000:0008:0800:200C:417A`
  - Shorthand rules can be applied:

    `2001:DB8:0:0:8:800:200C:417A`

    (Leading zeros within a 16-bit value can be removed)

    `2001:DB8::8:800:200C:417A`

    (One group of consecutive zeros replaced with double colon (:))

IPv6 is supported by the IBM XIV Storage System

IPv6 support is provided for XIV Management and VPN ports only and is supported for network functions such as DNS, SMTP, Lightweight Directory Access Protocol (LDAP), and Simple Network Management Protocol (SNMP).

Note: IPv6 address support is provided by both the XIV GUI and XCLI. For more details, see Chapter 7, “Monitoring” on page 327.

These specific functions do not support IPv6 (only IPv4 is used):

- Technician notebook port
- iSCSI port configuration
- Host iSCSI connections
- Mirroring iSCSI connections
- Support Center and XRSC communication

IBM XIV Storage System Software v11.1.1 and later offers US Government IPv6 (USGv6) compliance.

For more information about how to configure IPv6, see 5.3, “Configuring IPv6 addresses” on page 215.

3.2.3 Management connectivity

XIV Storage System is managed through three IP addresses over Ethernet interfaces on the patch panel to be resilient to two hardware failures. You must have three Ethernet ports available for management. If you require management to be resilient to a single network failure, be sure to connect these ports to two switches.

Make sure that the networking equipment providing the management communication is protected by a UPS.

Management IP configurations

For each of the three management ports, you must provide the following configuration information to the IBM SSR upon installation (see 3.2.1, “Basic configuration planning” on page 94):

- IP address of the port (all three ports must belong to the same subnet)
- Subnet mask
- Default IP gateway (if required)

Provide the following system-level IP information (not port-specific):

- IP address of the primary and secondary DNS servers
- IP address or DNS names of the SNMP manager, if required
- IP address or DNS names of the SMTP servers

Protocols

The XIV Storage System is managed through dedicated management ports running TCP/IP over Ethernet. Management is carried out through the following protocols (consider this design when configuring firewalls, other security protocols, and SMTP relaying):

- Proprietary XIV Storage System protocols are used to manage the XIV Storage System from the XIV Storage Management GUI and the XCLI. This management communication is performed over TCP port 7778, where the XIV Storage Management GUI/XCLI, as the client, always initiates the connection, and the XIV Storage System performs as the server.
- XIV Storage System sends and responds to SNMP management packets.
- XIV Storage System initiates SNMP packets when sending traps to SNMP managers.
- XIV Storage System initiates SMTP traffic when sending emails (for either event notification through email or for email-to-SMS gateways).
- XIV Storage System communicates with remote SSH connections over standard TCP port 22.

SMTP server

For correct operation of the Call Home function, the SMTP server must function as follows:

- Be reachable on port 25 for the XIV Storage System client-specified management IP addresses.
- Allow relaying from the XIV Storage System client-specified management IP addresses.
- Allow the XIV Storage System to send emails. The default sender address is `fromxiv@il.ibm.com`, but this address can be changed.

- Allow recipient addresses of `xiv-callhome-western-hemisphere@vnet.ibm.com` and `xiv-callhome-eastern-hemisphere@vnet.ibm.com`. Additionally, you can choose the shorter alias recipient address `xiv-west@vnet.ibm.com` and `xiv-east@vnet.ibm.com`.

3.2.4 IBM XIV Storage System physical installation

After all previous planning steps are completed and the system is delivered to its final location, the physical installation can begin. An IBM SSR completes all necessary tasks and the first logical configuration steps, up to the point where you can connect the IBM XIV Storage System through the XIV Storage Management GUI and the XCLI. Configuring storage pools, logical unit numbers (LUNs), and attaching the XIV Storage System to the host are storage administrator responsibilities. See Chapter 4, “IBM XIV Storage Management software” on page 111 for more information.

Physical installation

It is the responsibility of the client or moving contractor to unpack and move the XIV Storage System as close as possible to its final destination before an IBM SSR can start the physical installation. Carefully check and inspect the delivered crate and hardware for any visible damage. If there is no visible damage and the tilt and shock indicators show no problem, sign for the delivery.

Before starting the physical installation, ensure that an electrician is available who can fulfill the power requirements for connecting the XIV Storage System.

The physical installation steps are as follows:

1. Place and adjust the rack in its final position in the data center.
2. Check the XIV Storage System hardware. When the machine is delivered with the weight reduction feature (FC 0200), the IBM SSR installs the removed modules and components into the rack.
3. Connect the XIV Storage System power cords to the client-provided power source and advise an electrician to switch on the power connections.
4. Perform the initial power-on of the machine and perform necessary checks according to the particular power-on procedure.
5. To complete the physical steps of the installation, the IBM SSR performs various final checks of the hardware before continuing with the basic configuration.

Basic configuration

After the completion of the physical installation steps, the IBM SSR establishes a connection to the XIV Storage System through the patch panel (see 3.1.12, “Patch panel” on page 92) and completes the initial setup. You must provide the required completed information sheet that is referenced in 3.2.1, “Basic configuration planning” on page 94.

The basic configuration steps are as follows:

1. Set the Management IP addresses (client network), gateway, and netmask.
2. Set the system name.
3. Set the email sender address and SMTP server address.
4. Set the primary DNS and the secondary DNS.
5. Set the SNMP management server address.
6. Set the time zone.
7. Set the NTP server address.
8. Configure the system to send events to IBM (Call Home).
9. Configure and test remote support.

Completing the physical installation

After the IBM SSR completes the physical installation and initial setup, the IBM SSR performs the final checks for the XIV Storage System:

1. Power off and power on the machine using the XIV Storage Management GUI or XCLI.
2. Check the Events log carefully for problems.
3. Verify that all settings are correct and persistent.

The installation is complete, and the XIV Storage System is ready to be handed over to the client to configure and use. See Chapter 4, “IBM XIV Storage Management software” on page 111 for more information about that topic.

3.2.5 System power-on and power-off

Strictly follow these procedures to power on and power off your XIV Storage System.

Power on

To power on the system, complete these steps:

1. On each UPS, look for the Test button on the control panel (on the front of the UPS), as illustrated in Figure 3-34.

Figure 3-34 Locate Test button

Important: Do not confuse the Test button with the power-off button, which is normally protected by a cover. The Test button is the one circled in red in Figure 3-34.

2. Use both hands, as shown in Figure 3-35, to press each of the three Test buttons simultaneously.

Figure 3-35 Use both hands to press the three Test buttons simultaneously

This action starts applying power to the components in the rack and initiates the boot process for the interface modules and data modules.

Powering off the system

Powering off the system must be done only from either the XIV Storage Management GUI or the XCLI. You must be logged on as Storage Administrator (the storageadmin role).

Important: Do *not* power off the XIV Storage System by using the UPS power button, because this action can result in the loss of data and system configuration information.

Using the XIV Storage Management GUI

From the XIV Storage Management GUI, complete the following steps:

1. Click **Systems** → **System Settings** in the system main window toolbar and click **Shutdown System**, as shown in Figure 3-36.

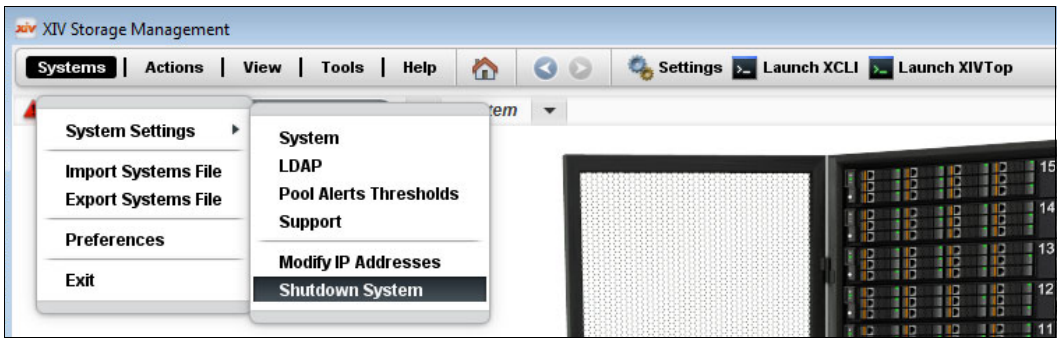


Figure 3-36 System shutdown

2. To confirm the shutdown, click **OK** and enter your login password, as shown in Figure 3-37.

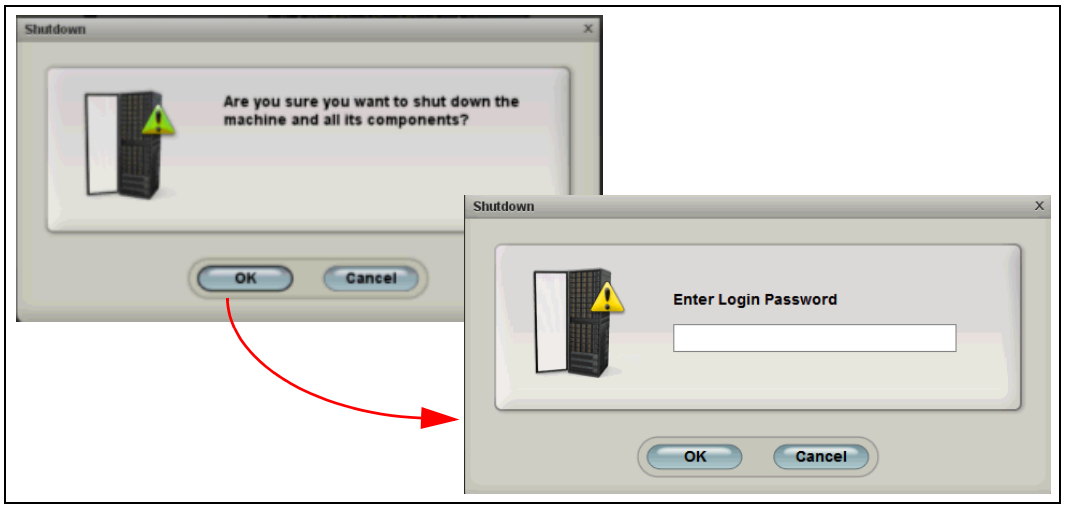


Figure 3-37 Confirm system shutdown

The shutdown takes less than 5 minutes. When it is finished, all fans and front lights on modules and all UPS lights are off.

Tip: Using the XIV Storage Management GUI is the most convenient and best way to power off the system.

Using the XCLI

From the command prompt, issue the following command, where XIV_01 is the system name:

```
xcli -c "XIV_01" -u admin -p adminadmin shutdown -y
```

You receive the response:

```
Command executed successfully
```

If you are using the XCLI session, use the shutdown procedure shown in Example 3-1.

Example 3-1 Executing a shutdown from the XCLI session

```
User Name: itso
Password: ********
Machine IP/Hostname: 9.155.53.250
connecting.
XIV_01>>shutdown
Password: ********

Warning:   ARE_YOU_SURE_YOU_WANT_TO_SHUT_DOWN y/n:
Command executed successfully
```

The shutdown takes less than 5 minutes. When it is finished, all fans and front lights on modules and all UPS lights are off.

Emergency power off

The XIV Storage System Machine Types 2810 and 2812 do not have an emergency power off (EPO) switch, but you can connect the XIV Storage System to a room emergency power off (EPO).

EPO precaution: Powering off the XIV Storage System using a room EPO switch results in data loss and possible loss of configuration. An IBM SSR is required to recover an XIV Storage System that was turned off using a room EPO switch. If the XIV Storage System loses ac power but is not powered off using an EPO circuit, data and configuration are preserved.

Local laws: National or local building, electrical, fire prevention, safety, and other laws or regulations can address or control the manner in which information technology equipment is installed within certain facilities and environments. The application of those laws or regulations can depend on considerations of factors beyond the nature or design of the equipment to be installed. It is a client's responsibility to interpret and identify any laws or regulations applicable to the installation of information technology in its environment and to inform IBM, IBM Business Partners, or their designated installers of any actions not identified in this document that are necessary to install information technology equipment in the client's facilities in accordance with such applicable laws or regulations.

Contact your IBM SSR for more information about connecting to a room EPO switch.

4

IBM XIV Storage Management software

This chapter describes the tasks performed by a storage administrator to configure the IBM XIV Storage System using the XIV Storage Management software. As of this publication, the current version of the XIV Storage Management software is Version 4.4.

The chapter covers the following topics:

- XIV Storage Management software introduction
- XIV Storage Management software installation
- XIV Storage Management software use
- Storage pools
- Volumes
- Host definition and mapping
- Multitenancy
- SSD caching
- Scripts

Important: Illustrations in this chapter apply mostly to an IBM XIV Storage System Gen3 fully configured with 2 or 3 TB disk drives.

The IBM Hyper-Scale Manager, which allows an administrator to monitor and configure many systems simultaneously, is discussed in the IBM Redpaper publication titled *IBM Hyper-Scale in XIV Storage,* REDP-5053.

© Copyright IBM Corp. 2014. All rights reserved.

4.1 XIV Storage Management software introduction

The XIV Storage Software runs internally on all system modules (both Data and Interface Modules) and provides all the functional capabilities. The functions and nature of this software are equivalent to what is typically referred to as *microcode* or *firmware* on other storage systems. Where on the XIV Storage System the firmware has a different meaning. Firmware within XIV means the vendor-specific, internal firmware of the dedicated hardware, such as the firmware of an InfiniBand host bus adapter (HBA) within an XIV module.

The XIV Storage *Management* software (also referred to as *XIV Management Tools*) is used to communicate with the XIV Storage System Software, which in turn interacts with the XIV Storage System hardware.

4.1.1 XIV Storage Management software platforms

The XIV Storage Management software can be installed on Microsoft Windows, Linux, AIX, HP-UX, Solaris, OSX, and iOS systems (available from the iOS App Store). The software serves as a management console for the XIV Storage System. The XIV Storage Management software is available for download at the following website:

`http://www.ibm.com/support/fixcentral/`

From this website, type 2810 in the Product selector and choose **XIV Storage System (2810, 2812)**. Then, choose the latest version under Installed Version and your operating system under Platform. Now, click **Continue**, as shown in Figure 4-1.

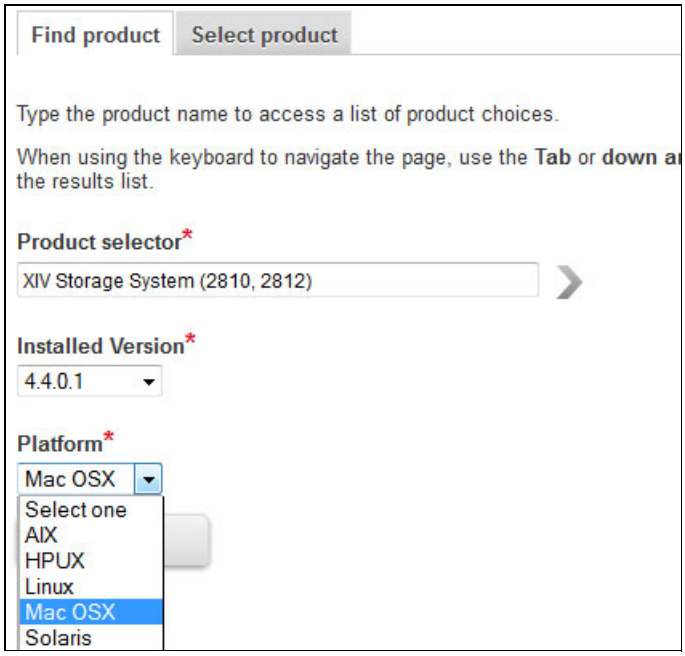


Figure 4-1 Fix Central portal

The Select fixes page is displayed. Scroll down to the section labeled Management tools and select **IBM_XIV_Management_Tools package**.

For more information about XIV Storage Management software compatibility, see the XIV interoperability matrix or the IBM System Storage Interoperation Center (SSIC):

`http://www.ibm.com/systems/support/storage/config/ssic/index.jsp`

4.1.2 XIV Storage Management software interfaces

The XIV Storage Management tools include a graphical user interface (XIV Storage Management GUI), an XIV Storage System Command-Line Interface (XCLI), and XIV Storage Management Top GUI (XIV Top), which is used to monitor performance.

XIV Storage Management GUI

The XIV Storage Management GUI (or XIV GUI) is an intuitive graphical user interface that lets you perform most administrative and technical operations (depending upon the user role) for one or several IBM XIV systems.

The motivation behind XIV Storage Management and the resulting GUI design is to eliminate the complexities of system management. Typical operational challenges, such as setup, configuration changes, general administration, and more, are achieved with a few clicks.

The XIV Management GUI supports two methods for connecting to IBM XIV systems:

- Direct mode to one or more IBM XIV systems, but acting on one XIV system at a time
- Manager mode, through the IBM Hyper-Scale Manager

> **Note:** The IBM Hyper-Scale Manager reduces operational complexity and enhances capacity planning through integrated management of multiple XIV systems. It is ideal for large and multi-site XIV deployments. For more information about the IBM Hyper-Scale Manager, see the IBM Redpaper, *IBM Hyper-Scale for the XIV Storage System*, REDP-5053.

This chapter contains descriptions and illustrations of tasks that are performed by a storage administrator when using the XIV Storage Management GUI in direct mode.

XIV Storage Management command-line interface

The XCLI is a powerful text-based tool that enables you to issue simple commands to configure, manage, or maintain the system, including commands used to connect to hosts and applications. The XCLI interface can be used in an XCLI Session environment to interactively configure the system or as part of a script to perform lengthy and complex tasks.

> **Tip:** The XIV Storage Management GUI executes XCLI commands during its operation. Therefore, anything that is available in the XIV Storage Management GUI can also be achieved through XCLI.

This chapter presents some of the common XCLI commands used by the administrator to interact with the system.

XIV Storage Management Top graphical user interface

The XIV Storage Management Top GUI provides the user with a simple tool for monitoring system performance on a host and volume level. See Chapter 7, “Monitoring” on page 327 for more information.

4.2 XIV Storage Management software installation

This section illustrates the step-by-step installation of the XIV Storage Management software under Microsoft Windows 2008.

GUI software: Although the illustrations in this chapter apply to an IBM XIV Storage System Gen3, the Version 4.4 XIV Storage Management GUI software also can be used with XIV Storage System second-generation systems.

Users running older versions are strongly encouraged to upgrade to Version 4.4 or later.

4.2.1 XIV Storage Management (GUI and XCLI) software upgrade

If a previous version of the XIV Storage Management GUI is installed, it is best to uninstall the existing version before installing the newer version.

The existing IBM XIV systems profile file, which is a list of defined IBM XIV systems and groups, is not deleted after the older version is uninstalled. This file is then recognized by the new installed version.

You can save the local IBM XIV systems profile file by exporting it from the XIV Storage Management GUI, as shown in Figure 4-2. It is saved as an XML file and is then available for importing into the newly installed XIV Storage Management GUI, which overwrites the existing systems profile.

Tip: The saved IBM XIV systems profile can be imported on extra systems on which the XIV Storage Management GUI is installed to ensure an identical systems view across all installations.

Figure 4-2 Export file

4.2.2 XIV Storage Management software installation steps for direct use

In this section, we describe the GUI installation procedure and its use for a *direct connection* to the XIV systems.

Important: The minimum requirements for installing the XIV Storage Management software in Windows 7 or Windows Server 2008 are as follows:

Processor:	Dual-core processor or equivalent
Memory:	512 MB of RAM (1 GB of RAM and above recommended)
Disk capacity:	150 MB
Screen resolution:	1024 x 768 (1024 x 768 - 1920 x 1200 recommended)

At the time of writing this book, we used an early version of the IBM XIV Storage Management GUI version 4.4. Later XIV Storage Management GUI releases might differ slightly in appearance.

Complete the following steps to install the XIV Storage Management software:

1. Locate the XIV Storage Manager installation file. Double-click the installation file and in the dialog window that is displayed, choose the language for the installation and click **OK**.

   Globalization: The XIV GUI Version 4.4 or later supports the English, Chinese, and Japanese languages

Figure 4-3 GUI install: GUI language

2. The initial installation window for the XIV Storage Management GUI installation is displayed. Click **Next** to continue.
3. The next window prompts you to accept the IBM License Software Agreement. Read the license, select **I accept both the IBM and the non-IBM terms** and click **Next** to proceed.

4. The installation window that is shown in Figure 4-4 is where the installation directory is specified. Enter an installation location or proceed with the default location provided.

   Click **Next** to proceed.

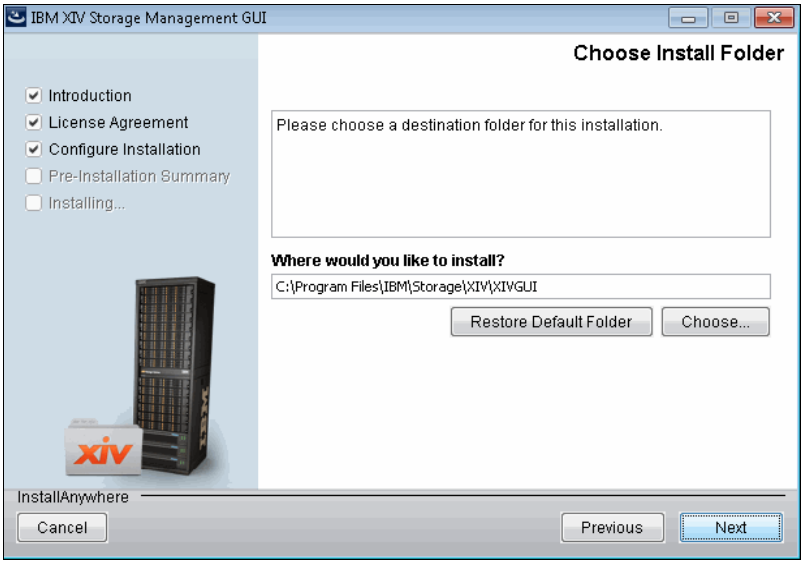


Figure 4-4 Choose the installation directory

5. For new installations, or if installing in a new directory and a previous GUI installation exists, a program shortcut called IBM XIV 4.4 is created in the Start menu folder. In addition, three optional desktop icons are created, as indicated in Figure 4-5. If you choose to install into a folder that already contains an installation of the GUI, this dialog is skipped. Therefore, click **Next** to proceed.

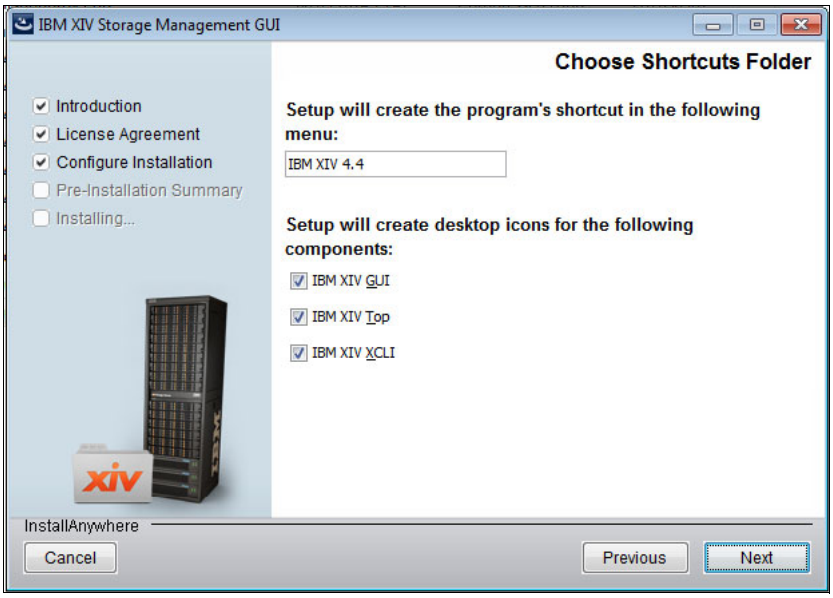


Figure 4-5 Select Start menu folder and desktop icons

6. The setup configuration confirmation window is now displayed. Review your selections and click **Install** to complete the installation.
7. Finally, the Install Complete window shown in Figure 4-6. Click **Done** to exit the installer.

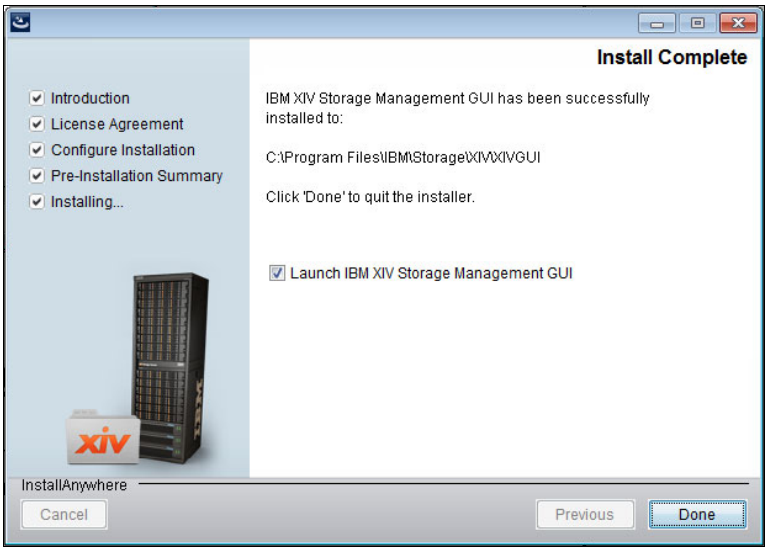


Figure 4-6 Completing the setup

4.3 XIV Storage Management software use

Now that the XIV Storage Management software has been installed, the introduction and explanation of some features of the XIV Storage Management GUI and XCLI can be useful in further configuration and storage management.

Beginning with the XIV Storage Management GUI, this section describes the following topics:

- Signing into the XIV Storage Management GUI
- Connecting to IBM XIV systems
- Overview of the management and system views
- Review of XIV Storage Management GUI features

Next, moving to the XCLI, the following topics are described:

- Launching the XCLI
- XCLI Session features
- Customizing the XCLI environment
- Using XCLI help

4.3.1 XIV Storage Management GUI used in direct mode

The XIV Storage Management GUI is a multi-function tool for individually managing one or more IBM XIV systems. This section describes the basic steps to follow to get started using the GUI.

Signing onto the XIV Storage Management GUI

When starting the XIV Storage Management GUI for the first time, a login window prompts you for a user name and its corresponding password before granting access to the XIV Storage System. The default user is *admin* and the default corresponding password is *adminadmin* (see Figure 4-7).

Storage Management
User:
admin
Password:
Demo Mode
Login
Licensed Materials - Property of IBM Corporation and other(s). Copyright © 2008, 2012. IBM and XIV are registered trademarks of IBM Corporation in the United States, other countries, or both.

Figure 4-7 Login window with default access

Important: Remember to change the default passwords to correctly secure your system. For more information, see 5.5.5, “Password management and resets” on page 237.

The default admin user comes with the storage administrator (*storageadmin*) role. The XIV Storage System offers role-based user access management that consists of the already mentioned storage administrator, application administrator, security administrator, and read-only role.

The direct mode access is allowed for stand-alone instances of the GUI on users, workstations, and desktops, as well as the demo mode.

For more information about user security and roles, and how to manage multiple storage systems, see Chapter 5, “Security” on page 207.

Tip: The Demo Mode option, which is seen in Figure 4-7, is accessed by selecting the **Demo Mode** check box and clicking **Login**. No credentials are required. This demo mode is useful for learning how the XIV Storage Management software works without needing an actual XIV system.

Note: If the GUI is configured to connect to an IBM Hyper-Scale Manager, the login dialog window shows an additional Mode drop-down field. The following drop-down options are available:

- Direct, for direct mode connection
- Manager, for connecting via an IBM Hyper-Scale Manager
- Demo, for running the GUI in demonstration mode

For more information about using Manager mode, see the IBM Redpaper titled *IBM Hyper-Scale in XIV Storage*, REDP-5053.

Connecting to an XIV Storage System

To connect to an XIV Storage System, initially add the system to make it visible in the XIV Storage Management GUI by specifying its IP addresses.

To add the XIV Storage System, complete the following steps:

1. Ensure that the management workstation is set up to access the LAN subnet where the XIV Storage System is located.

   Tip: It is a good idea to verify the connection by pinging the mangement IP addresses of the XIV Storage System.

   When starting the XIV Storage Management GUI on the management workstation for the first time, the Add System Management window automatically opens.

2. Enter the IP address (or set of IP addresses, for redundancy) of the XIV Storage System in the IP/Host name fields. Click **Add** to add the system to the XIV Storage Management GUI. See Figure 4-8.

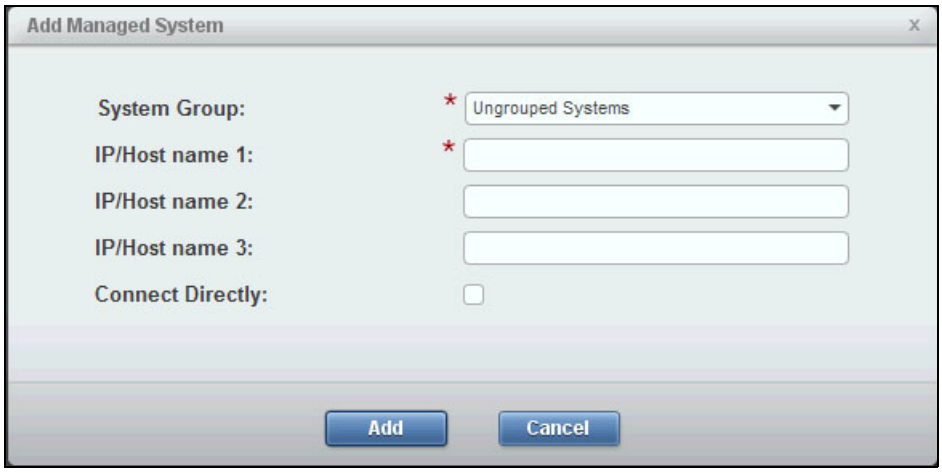


Figure 4-8 Add Managed System window

Tips: XIV System software 11.1 and later support IPv6 addressing for the XIV Storage System management ports. For information about configuring IPv6, see 5.3, “Configuring IPv6 addresses” on page 215.

If the workstation is correctly configured for Domain Name Service (DNS) resolution, host names can be used.

Important: The Connect Directly check box is for an IBM service representative to use to connect their notebook to the XIV through the two notebook ports on the patch panel. This option must not be used by general GUI users.

3. The XIV Storage Management system view window opens, as shown in Figure 4-9. Further administrative actions can be taken.

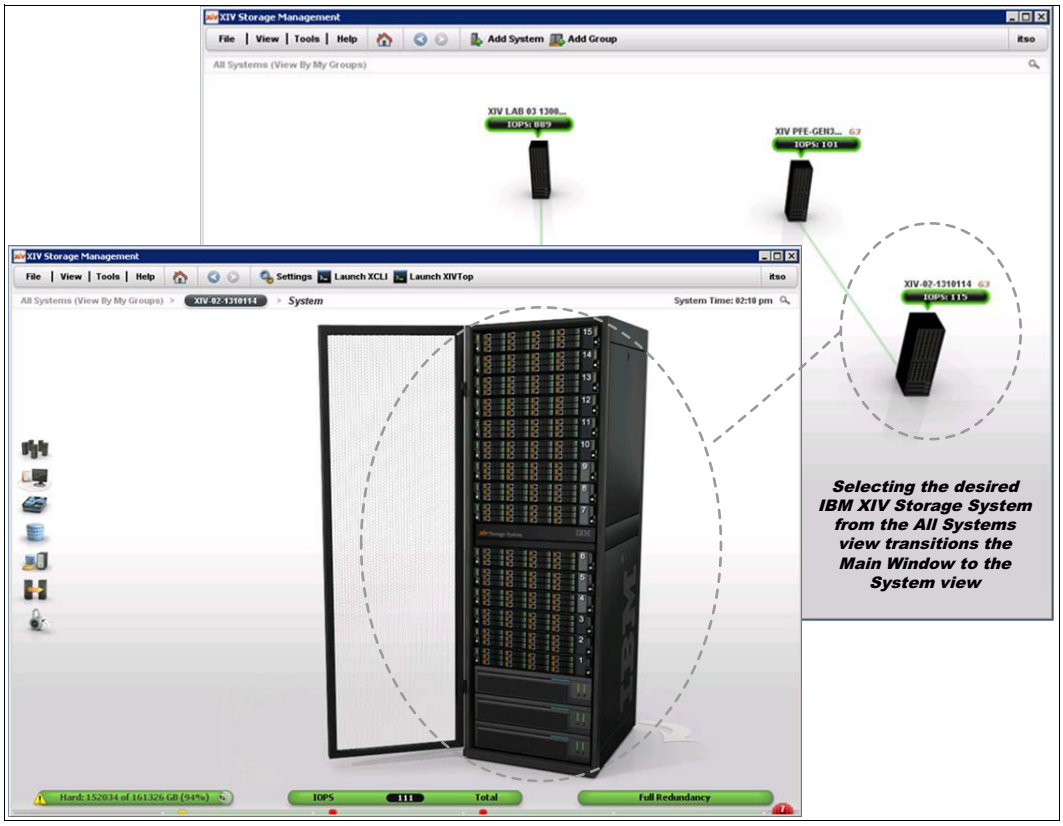


Figure 4-9 XIV foreground: Storage Manager main window, System view. XIV background: All Systems view

4. If multiple IBM XIV systems are configured for management in the GUI, each system is always visible in the All Systems view, regardless of the credentials provided.

   However, only systems that are successfully authenticated with the current credentials are accessible. Move the cursor over the appropriate XIV Storage System and click it to open the XIV Storage System Management view of that system. To return to the All Systems view, click the back arrow in the menu bar, the All Systems link in the hierarchical navigation or the home icon, as shown in Figure 4-10.

Figure 4-10 Navigate back to All Systems view

5. For ease of management, the GUI supports logical grouping of IBM XIV systems:
   a. To create a group, right-click an empty area of the All Systems view and select **Add Group**, as shown in Figure 4-11.

Figure 4-11 Add Group

> **Tip:** To create a group, there must be more than one XIV Storage System configured in the GUI.

   b. The Add New Group dialog that is shown in Figure 4-12 is displayed. Enter a name for the new group and click **OK**.

Figure 4-12 Add New Group

   c. The new group now is displayed in the All Systems view. See Figure 4-13 on page 122.

   You can add a maximum of 12 storage systems to a group and you can have a maximum of 12 groups.

   Systems already configured in the GUI that are not part of a group can be added to a group by dragging and dropping them onto the wanted group.

   To add a system to a group, right-click the **Group** and select **Add System** (Figure 4-14 on page 123). In this case, the Add Managed System dialog box is shown as in Figure 4-8 on page 119. Complete the dialog box as explained in 4.3.1, “XIV Storage Management GUI used in direct mode” on page 117.

Figure 4-13 shows the XIV Group on the All Systems view.

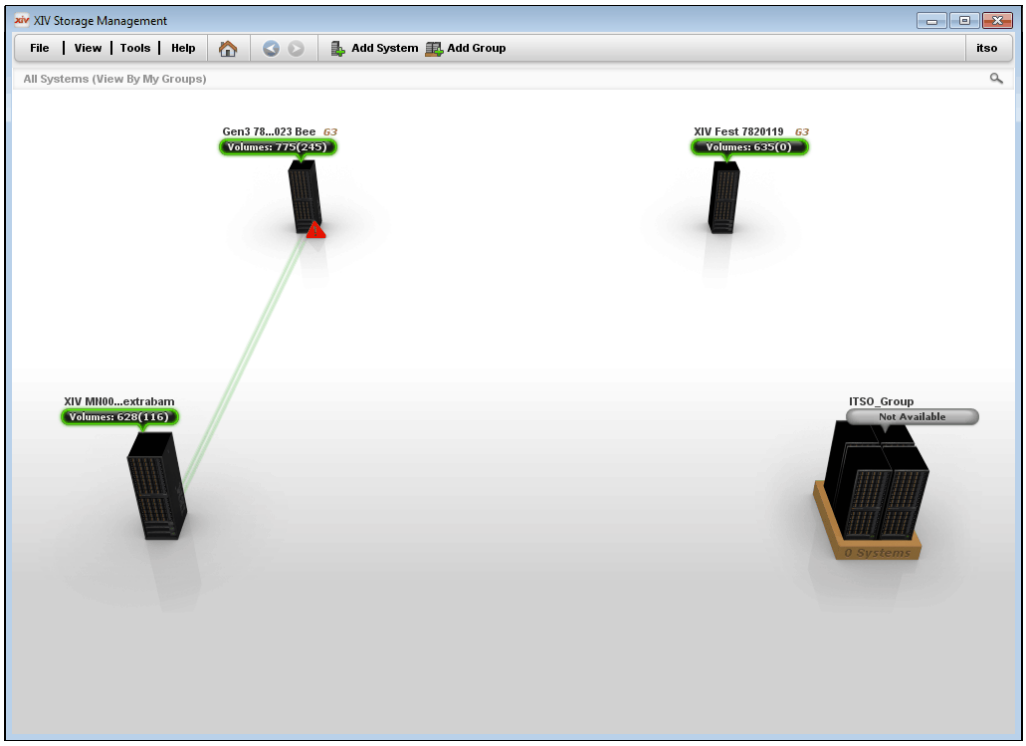


Figure 4-13 XIV Group on All Systems view

Tip: With the XIV GUI v4.4 in direct mode, you can manage up to 12 groups in the All Systems view. A group can contain a maximum of 12 IBM XIV systems, and any particular system can belong to only one group. A single GUI instance can manage a maximum of 144 systems. To manage more than 144 systems in direct mode, multiple instances of the GUI can be started, each with a different profile. For more information, see “Direct mode GUI limitations for managing multiple systems” on page 240. However, it is suggested to use the IBM Hyper-Scale Management server when managing more than 40 IBM XIV systems. For more information, see the IBM Redpaper, *IBM Hyper-Scale for the XIV Storage System*, REDP-5053.

d. Additional actions are available to delete or rename existing groups. A menu is displayed by right-clicking a group, as shown in Figure 4-14 on page 123.

- If you select **Rename Group**, a dialog box is displayed and you can enter a new group name.
- If you select **Delete Group**, the group is removed.

Note: Deleting a group that contains systems also removes those systems from the GUI. If you want to delete a group without removing the systems that it contains, first move the systems out of the group by right-clicking each system and choosing the **Move System to** option.

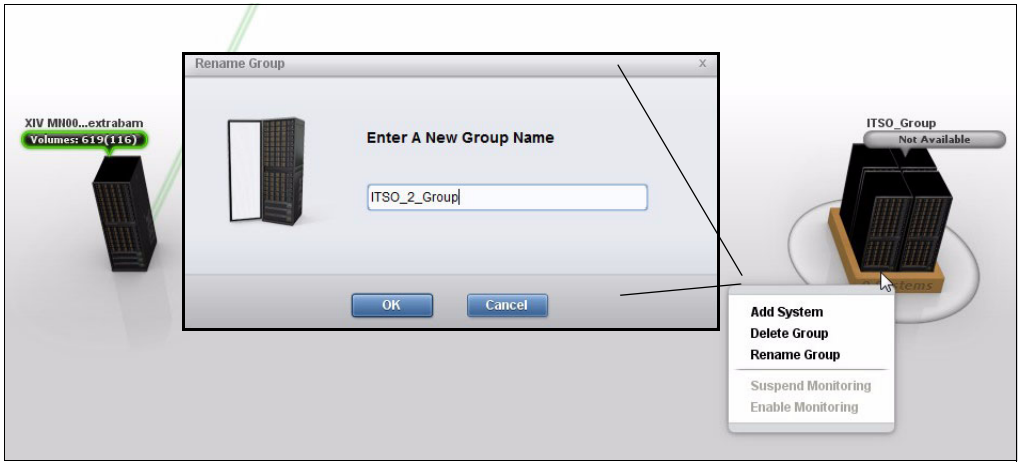


Figure 4-14 Delete and Rename Group options

XIV Storage Management software features

Both the XIV Storage Management GUI and XCLI are mostly self-explanatory, with organized structures and simple navigation. Some of the standard and newer unique features are described in further detail in this section.

XIV Storage Management GUI main system view

Management of the XIV Storage System begins at the main System view, which is shown in Figure 4-15. From this view, the administrator can quickly get the status of the system. Depending on the user permissions, the interface might differ slightly.

Figure 4-15 Storage Manager main window: System view

Perspective: The figures listed in this chapter depict the interface layout from a storage administrator perspective. Not all options are available to the application administrator and read-only permission levels.

The main window, which is the System view, is divided into the following areas, as shown in Figure 4-15 on page 123:

Function icons and menus — On the left side of the main window in the System view, a set of vertically stacked icons is used to navigate between the functions of the XIV Storage Management GUI. Move the cursor over an icon to preview the function menu for that icon. A brief description of each of the function icons is shown in Figure 4-16 on page 125.

Main display — This provides a graphical representation of the XIV Storage System. Move the cursor over a specific hardware component (module, disk, uninterruptible power supply [UPS] unit) to open a status callout for that component. Click the arrow at the lower right to see a view of the system patch panel. Move the cursor over a port to open a status callout for that port.

Menu bar — This area is used for configuring the system and as an alternative to the function icons. Also, various user tools pertaining to XIV Storage System are available here.

Toolbar — This area contains a dynamic, contextual set of actions based on the current view.

User indicator — Identifies the current authenticated user. Click the user indicator to log in with alternate credentials.

System information menu — Displays details about the current system. To change the displayed details, click the small triangle menu at the left of this menu. The following details can be displayed:

- Hardware details
- System Time: The time on the XIV system
- Local Time: Time on the management workstation
- System software version

Hierarchical navigation — Provides a hierarchical view of the current system or group, with contextual drop-down menus.

Current systems indicator — Indicates the currently selected systems or group.

Status bar and Status area — These indicators are at the bottom of the main window of the System view. This area indicates the overall operational status of the currently selected IBM XIV Storage System or group of systems. The values displayed are in the context of the current systems:

The *left* indicator shows the amount of soft or hard storage capacity currently allocated to the storage pools and provides alerts when certain capacity thresholds are reached. Click the indicator to toggle between soft and hard values. As the physical, or hard, capacity consumed by volumes with the storage pool passes thresholds, the

color of this meter indicates that extra hard capacity might need to be added to one or more storage pools.

The *center* indicator displays the current number of I/O operations per second (IOPS).

The *right* indictor shows the general system status and various operational states of the XIV Storage System. An alerts icon might be displayed that, if selected, shows alerts and corresponding events. For more information, see 7.1, “Monitoring with XIV Storage Management GUI and XCLI” on page 328.Figure 4-16 shows the XIV Storage Management software menu items.

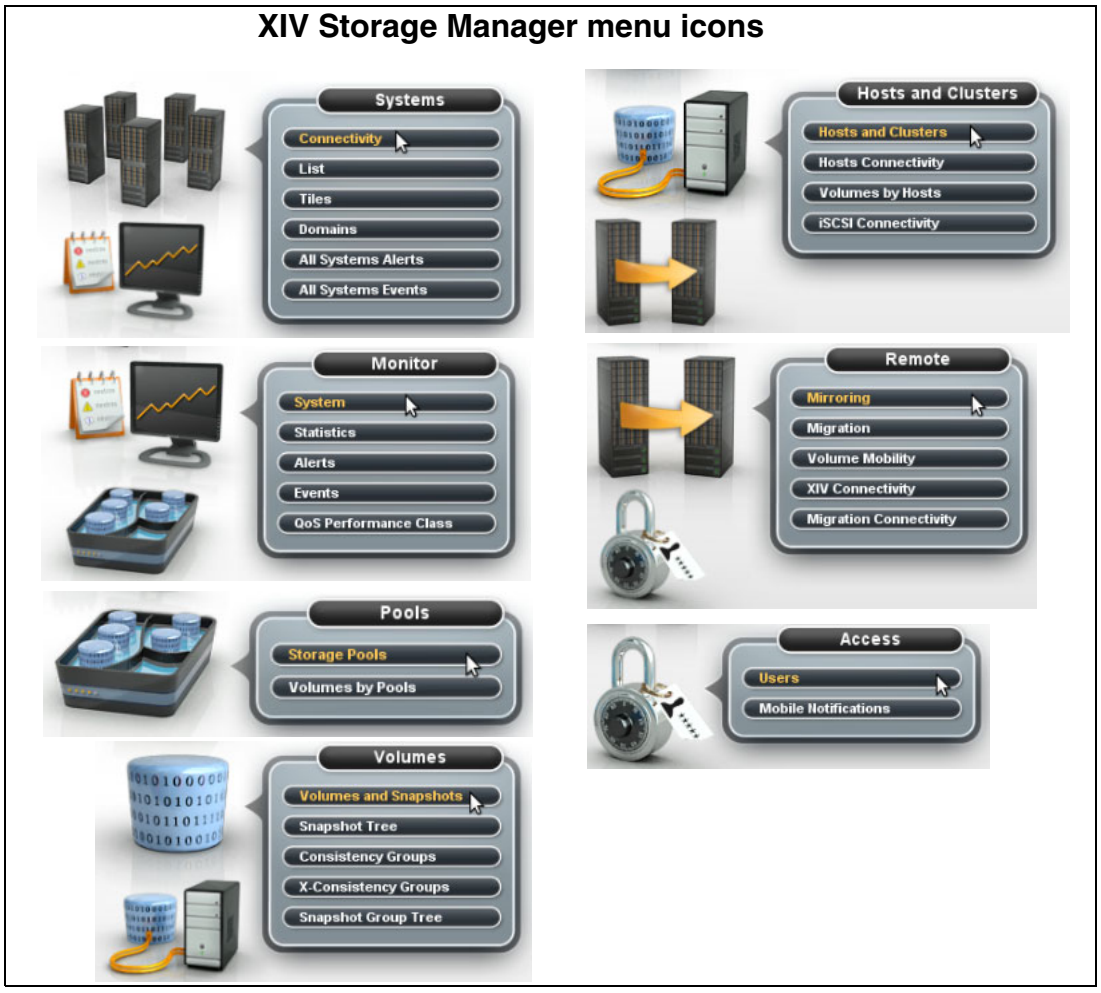


Figure 4-16 Menu items in XIV Storage Management software

Tip: Depending on the current user role, more or fewer systems administrative actions might be available in the menu items.

XIV Storage Management GUI features

Several notable features in XIV Storage Management Version 4.4 facilitate efficient system configuration and management by using the XIV Storage Management GUI.

Tree Table view

When working with the XIV Storage Management GUI, a standardized view has been introduced called the *Tree Table* (TT) view. As shown in Figure 4-17 in the *Pools - Volumes by Pools* view, this feature presents the navigation of several views (Pools, Volumes, Hosts, and Clusters) using a tree-like table view. This view can be useful when accessing the current storage-related views and making changes accordingly.

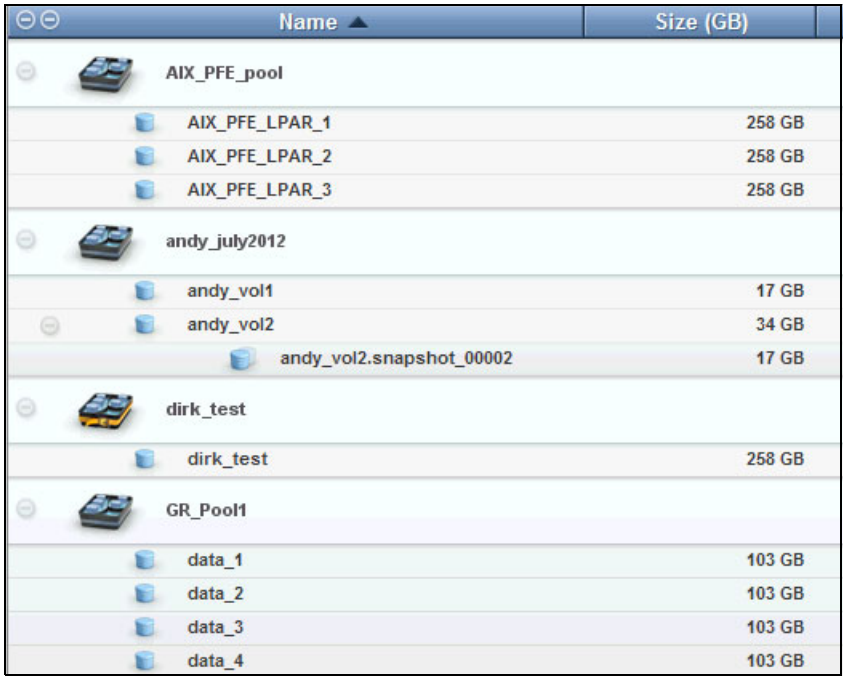


Figure 4-17 Tree Table

Export to comma-separated values format

In addition to the previously supported events and statistics frames, the XIV Storage Management GUI Version 4.4 supports exporting to comma-separated values (CSV) for *all* GUI tabular views. The new Export icon is available for all of these views, as shown in Figure 4-18. When you click the icon, a window opens to prompt you to save the export file to your local workstation.

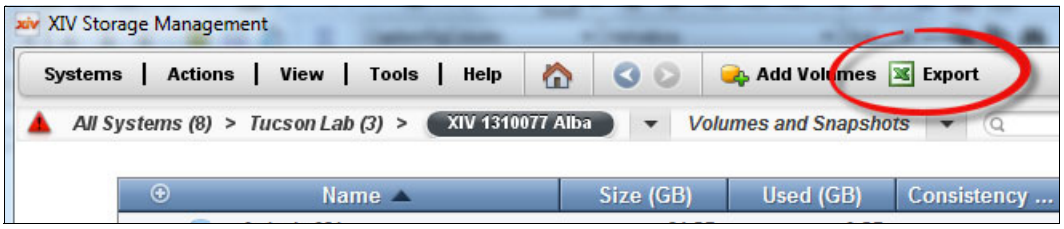


Figure 4-18 Export to CSV icon

System management: General settings

To reach the general settings window, use the GUI **Tools** menu (Alt + t), and then click **Management**. A new option introduced in GUI 4.4 is the GUI session timeout setting, as highlighted in Figure 4-19. It is a security improvement, introduced for conformance to the Payment Card Industry Data Security Standards (PCI DSS). The default setting is 15 minutes and can be adjusted by the user. This setting means that GUI and XCLI will become locked after the defined idle timeout setting and for further activity, the user must reauthenticate.

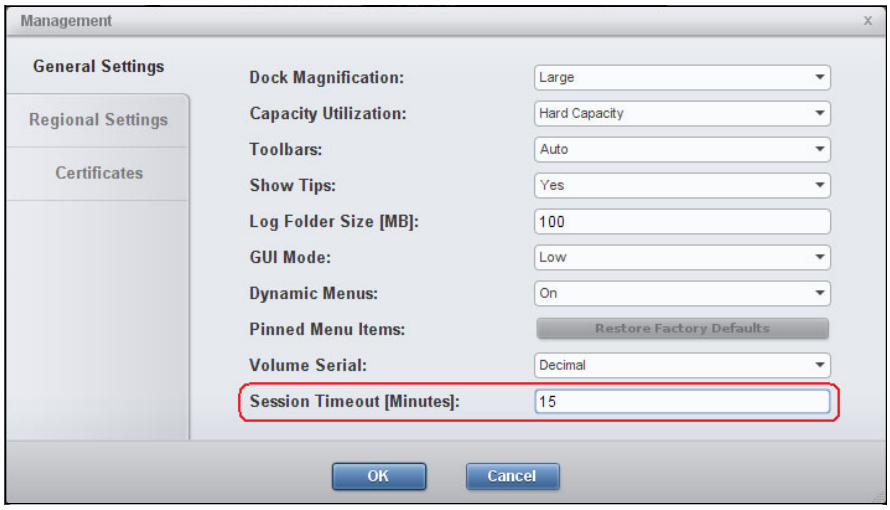


Figure 4-19 General settings in the GUI: GUI session timeout

Regional settings and time zone awareness

It is common for storage administrators to manage storage systems from locations that are geographically separated from those systems. GUI Version 4 enhances support for this scenario by introducing time zone awareness. Throughout the GUI, many details can be displayed in both system time (time on the XIV Storage System) and local time (time on the workstation running the GUI). To configure the regional settings for the GUI, click **Tools** in the menu bar, click **Management**, and then click the **Regional Settings** tab (see Figure 4-20).

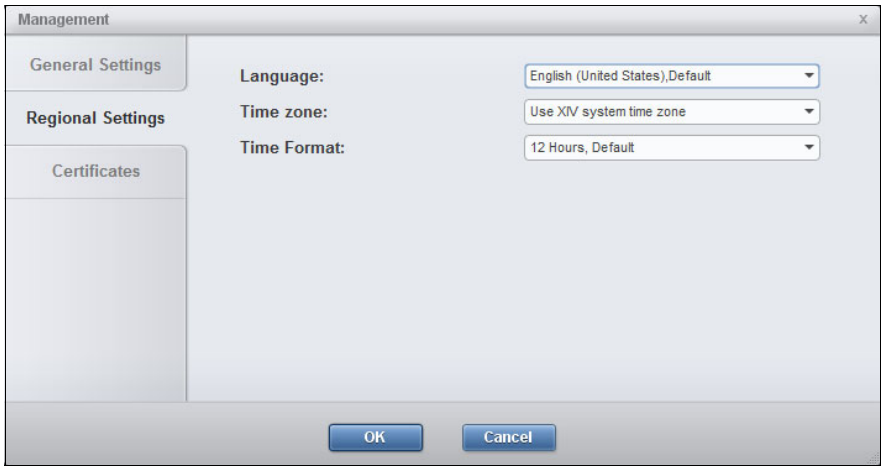


Figure 4-20 Regional Settings tab in the GUI

Disabled menu item tooltips

For each GUI menu item that is disabled, the menu item contains a tooltip that explains the reason why the item is disabled. To view the tool tip, move your cursor over the circled ***i*** icon at the right of the menu item, as shown in Figure 4-21.

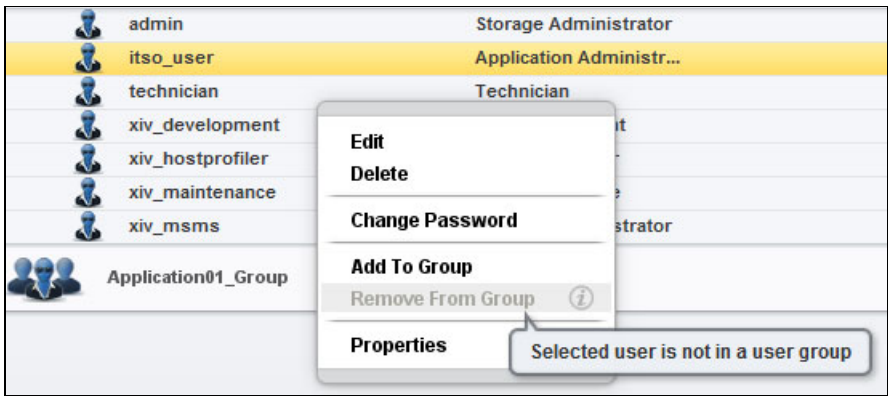


Figure 4-21 Disable menu item tooltips

Secure Sockets Layer certificate management

Enhanced Secure Sockets Layer (SSL) certificate support in XIV System Management software 4.4 improves the flexibility and security of managing IBM XIV systems in two ways:

- Support for the use of self-signed certificates for securing communications between the GUI and XIV Systems running System software Version 11.2 or later, whether via direct mode, or via the MultiSystem Manager. To manage the self-signed certificates on an XIV Storage System, click **Systems** in the menu bar, click **System Settings**, then click **Manage Certificates**, as shown in Figure 4-22. For more details about certificate management, see 5.2, “x509 certificate validation and management” on page 208.

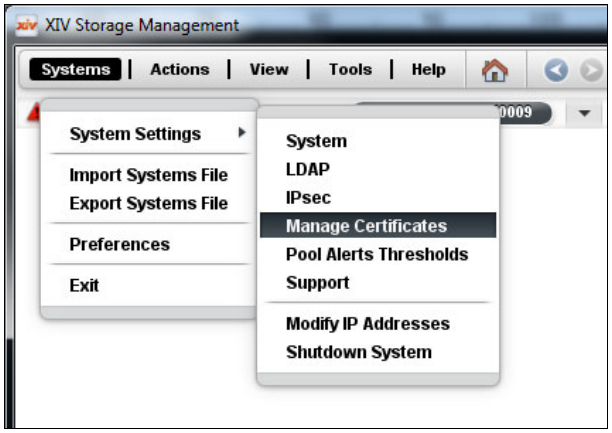


Figure 4-22 Manage XIV system certificates

Note: The **Manage Certificates** menu option is not visible for XIV Systems running system software older than Version 11.2.

- In addition, the GUI now supports the management of the trusted certificates used to secure its communication with IBM XIV systems. To facilitate the import and management of trusted certificates in the GUI, a new Certificates tab has been added to the Management view. To access, click the **Tools** portion of the menu bar, click **Management**, and then click the **Certificates** tab in the Management window, as shown in Figure 4-23. For more information about the use of certificates, see 5.2, “x509 certificate validation and management” on page 208.

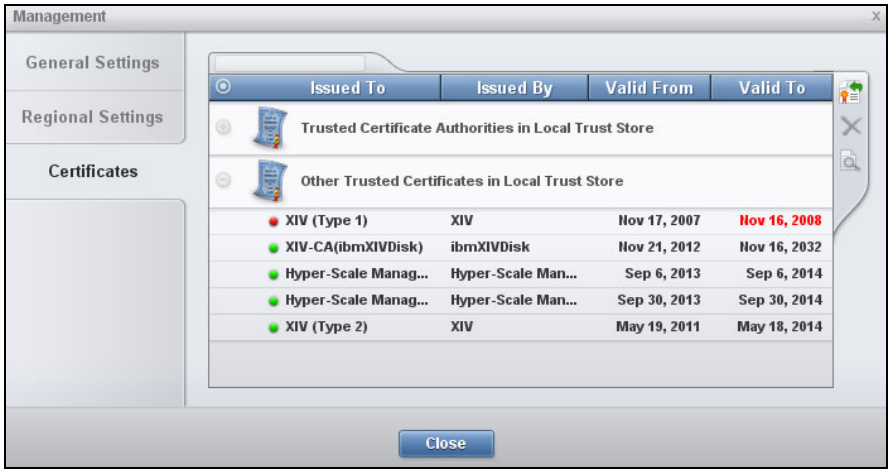

Management
General Settings
Regional Settings
Certificates
Issued To
Issued By
Valid From
Valid To
Trusted Certificate Authorities in Local Trust Store
Other Trusted Certificates in Local Trust Store
XIV (Type 1) XIV Nov 17, 2007 Nov 16, 2008
XIV-CA(ibmXIVDisk) ibmXIVDisk Nov 21, 2012 Nov 16, 2032
Hyper-Scale Manag... Hyper-Scale Man... Sep 6, 2013 Sep 6, 2014
Hyper-Scale Manag... Hyper-Scale Man... Sep 30, 2013 Sep 30, 2014
XIV (Type 2) XIV May 19, 2011 May 18, 2014
Close

Figure 4-23 Certificate management in the GUI

Copy and paste of system configuration settings

To facilitate efficient configuration of system settings when managing multiple IBM XIV systems, the GUI supports copying the settings of one system to paste those settings to one or more other systems.

To use this feature, first right-click the system that contains the configuration settings that you want to copy and choose **Copy System Configuration**, as shown in Figure 4-24 on page 130.

Next, right-click a target system and choose to paste the appropriate settings. The following configuration settings can be pasted to the target system, each in a separate step:

- Support Configuration
- LDAP Configuration
- Pools Alerts Thresholds

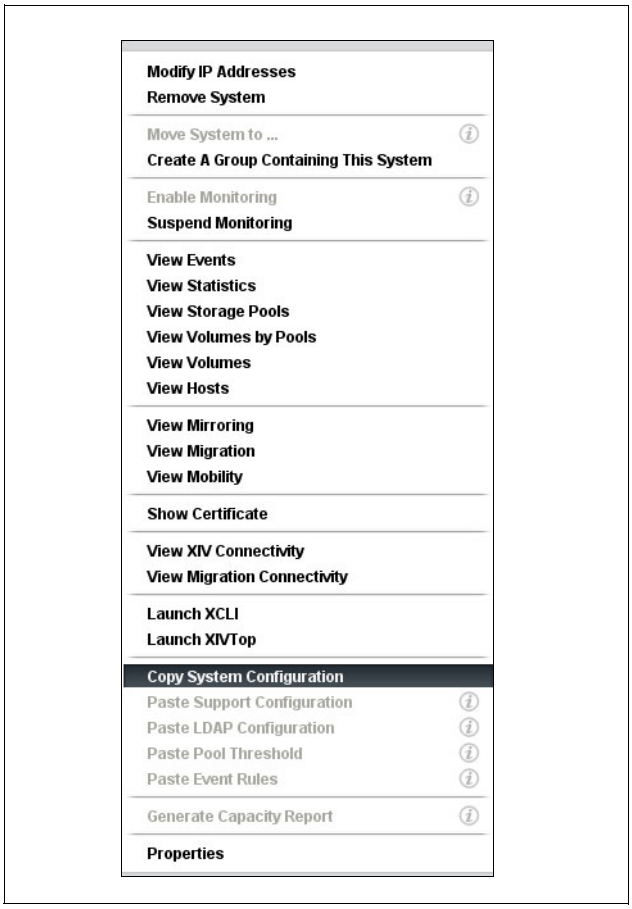


Figure 4-24 Copy and paste system configuration settings

After you select which settings to paste, a dialog window is displayed, as shown in Figure 4-25. This window confirms your selection and allows you to begin copying the settings to the target system, or to cancel the operation. If you choose to proceed, the dialog window shows the progress of the operation as it completes.

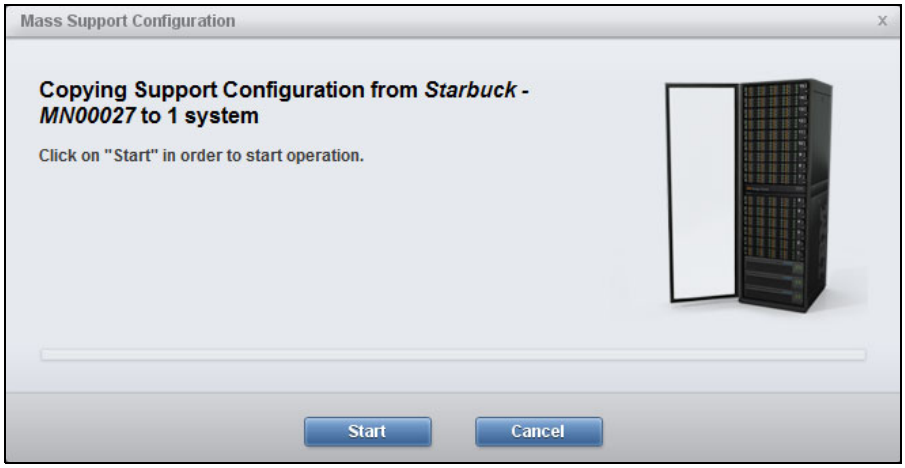

Mass Support Configuration
Copying Support Configuration from Starbuck - MN00027 to 1 system
Click on "Start" in order to start operation.
Start
Cancel

Figure 4-25 Copying configuration settings: Dialog window

System Selector menu

Available from the hierarchical navigation is a quick access System Selector menu for viewing the status of all systems under management and easy navigating to a specific system. This menu is available from the All Systems and group views. To access the System Selector menu, press Alt-S or click the triangle icon to the right of the All Systems or Group Name links, as shown in Figure 4-26. Press Alt-S repeatedly to cycle though the systems that are shown in the menu.

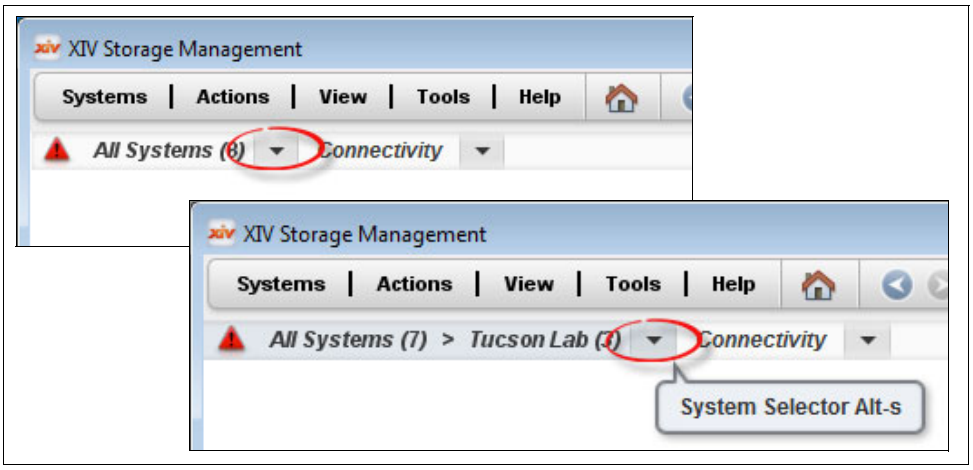

XIV Storage Management
Systems | Actions | View | Tools | Help
All Systems (8)
Connectivity
XIV Storage Management
Systems | Actions | View | Tools | Help
All Systems (7) > Tucson Lab (3)
Connectivity
System Selector Alt-s

Figure 4-26 Access the System Selector menu

When activated, the System Selector menu is displayed, as illustrated in Figure 4-27.

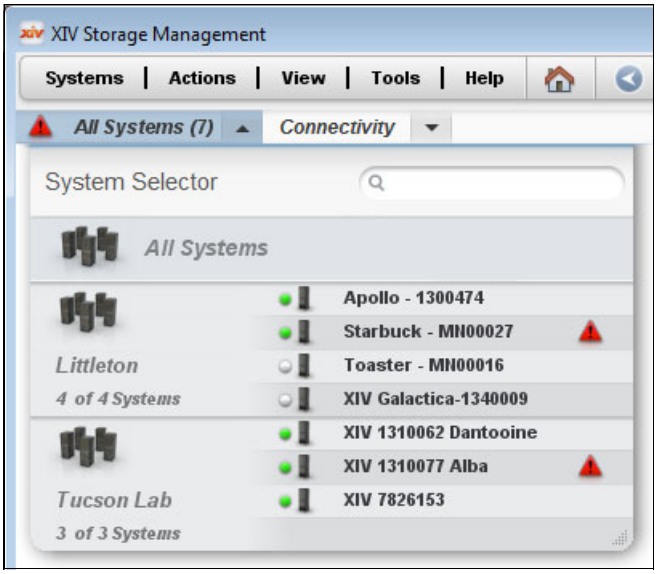


Figure 4-27 The System Selector menu

Tip: To select multiple systems from the System Selector menu, hold the Ctrl key while clicking each system.

Views Selector menu

The GUI supports several view styles for displaying information about the systems being managed. To change the current view, use the Views Selector menu by pressing Alt-X, or by

clicking the triangle icon at the far right of the hierarchical navigation, as shown in Figure 4-28.

Figure 4-28 Views Selector menu

When viewing all systems or a group of systems, three useful views are available. Use the Views Selector menu to modify the current view to best suit the management task you are performing.

The Connectivity view, shown in Figure 4-29 on page 132, provides a three-dimensional view of the selected systems, with visual indicators for replication relationships between systems and group membership. For each system, the system name, serial number, system software version, and hardware version (Gen3 only) are displayed. Click the balloon above a system to change the balloon value. The following values are available:

- System IOPS
- System Utilization
- System Status
- System Number of Hosts
- System Number of Volumes
- System Hardware Type

Figure 4-29 All Systems Connectivity view

Note: When one or more systems are selected in the Connectivity view, the cleared systems appear grayed out.

The Tiles view, shown in Figure 4-30, displays the selected systems in a grid format. For each system, the following details are displayed:

- Group name (if applicable)
- System name and serial number
- System software version
- Current system IOPS

- Capacity use
- System status, including indicators for hardware problems or alerts

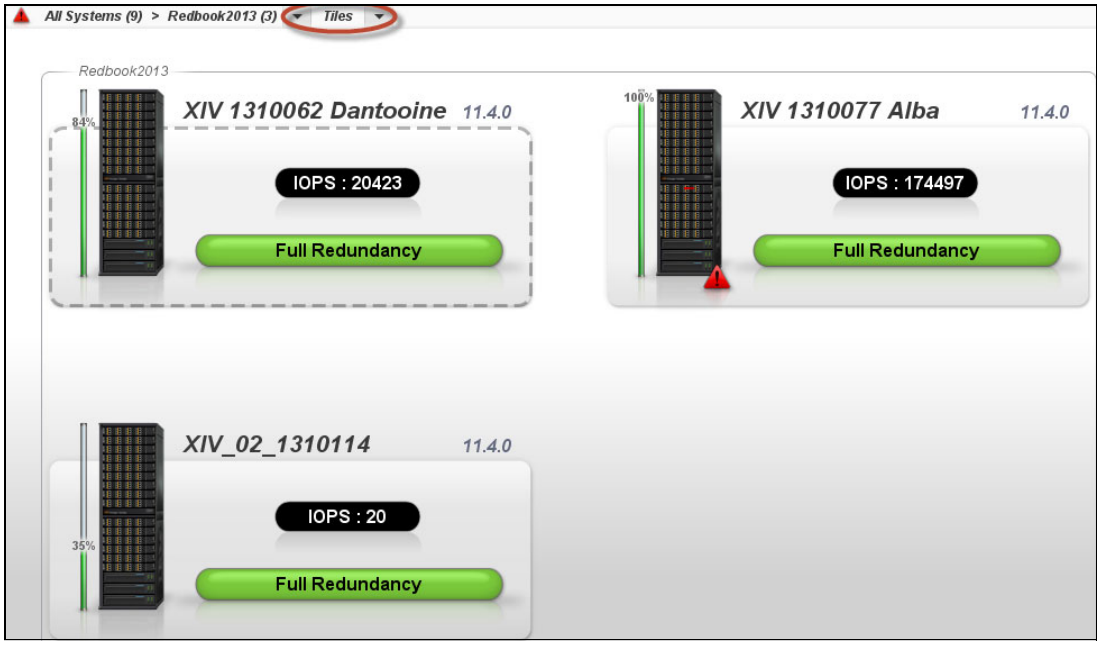


Figure 4-30 All Systems, Tiles view

Note: Only selected systems are visible in the Tiles view. Cleared systems are hidden.

Tip: To show or hide systems in the Tiles view, use the System Selector menu, as shown in Figure 4-27 on page 131.

The List view, shown in Figure 4-31, displays the details of the selected systems in tabular format. The listed systems can be sorted by any column. Click a column header to sort by that column. Click that header again to toggle between ascending and descending order. Right-click a column header to customize which column header values are shown or hidden. The Name, Group, and Status columns are required. The following optional columns are available:

- IOPS
- Hard Size
- Hard Used
- Hardware Type
- Hosts
- Soft Size
- Soft Used
- Utilization
- Version
- Volumes

Double-click any system row to navigate to the System view for that system.

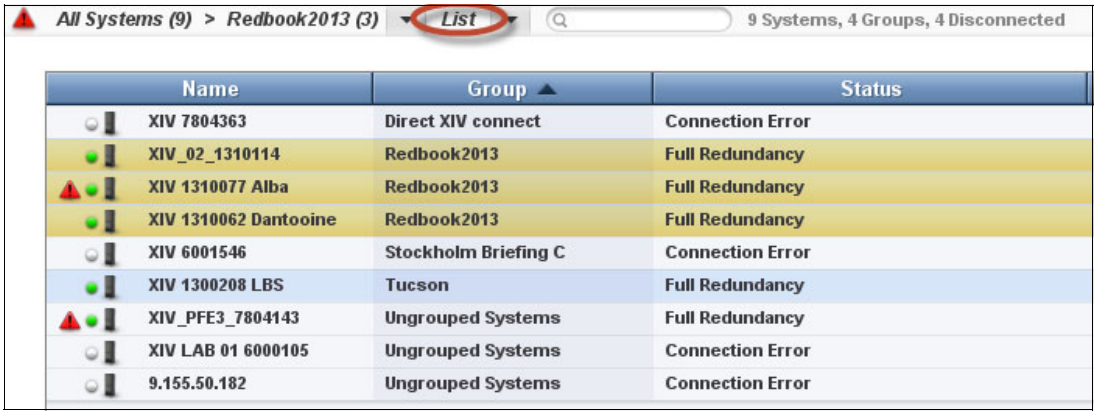


Figure 4-31 All Systems List view

Note: When one or more systems are selected in the List view, the selected rows are highlighted in orange.

Tip: From any of the Systems views, you can select multiple systems by holding the Ctrl key while clicking each system. While one or more systems are selected, you can perform an action on those systems together by right-clicking any selected system. This action brings up a contextual menu, allowing you to perform the wanted action.

Messages

When the Status area alerts the user, messages of real-time problems in the environment are reported. See Figure 4-32 for an example of a real-time alert that is displayed at the upper-left side of the GUI panel. These are examples of the types of messages:

- Hardware problem (disk, modules, fan, or service)
- Utilization problem (pool hard capacity)
- System status
- Uncleared alerting events
- System-to-system connection problem

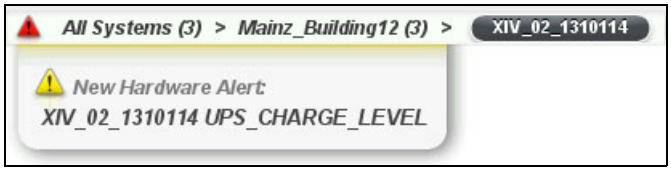


Figure 4-32 Event messaging

If a configuration-related message is noted and clicked, regardless of which XIV Storage System is being administered, the user is taken directly to the system in question to further review the message contents. The user is taken to the event log where the additional details can be viewed. See “All Systems Events window” on page 329.

Search and view filtering

The XIV Storage Management GUI features a Search function. This function allows for easy searches within the XIV Storage Management GUI for configuration-related items. You can search across all the GUI managed storage systems, and you can search either by name only or by object type.

To start the search function, click the magnifying glass in the upper right corner of any of the XIV Storage Management GUI windows, as illustrated in Figure 4-33. Alternatively, you can press Ctrl-F to start the search function.

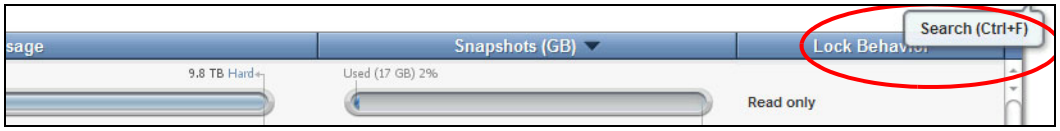


Figure 4-33 Search in the GUI

A Search panel opens as illustrated in Figure 4-34.

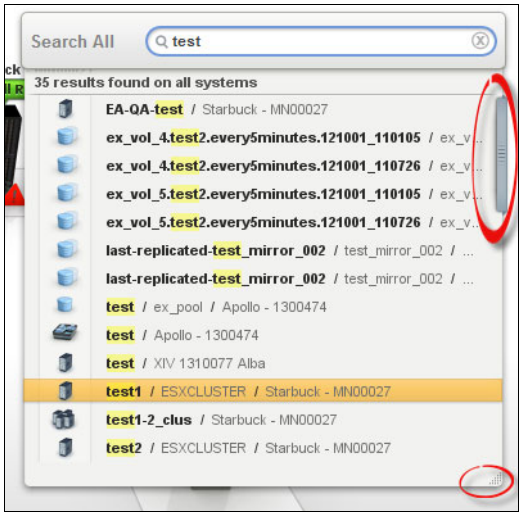

Search All
test
35 results found on all systems
EA-QA-test / Starbuck - MN00027
ex_vol_4.test2.every5minutes.121001_110105 / ex_v...
ex_vol_4.test2.every5minutes.121001_110726 / ex_v...
ex_vol_5.test2.every5minutes.121001_110105 / ex_v...
ex_vol_5.test2.every5minutes.121001_110726 / ex_v...
last-replicated-test_mirror_002 / test_mirror_002 / ...
last-replicated-test_mirror_002 / test_mirror_002 / ...
test / ex_pool / Apollo - 1300474
test / Apollo - 1300474
test / XIV 1310077 Alba
test1 / ESXCLUSTER / Starbuck - MN00027
test1-2_clus / Starbuck - MN00027
test2 / ESXCLUSTER / Starbuck - MN00027

Figure 4-34 Search window

Enter search text in the field to perform a search across all of the XIV systems that are managed by the GUI. Results are displayed in a Tree Table view. Clicking an item in the search results navigates to that item in the GUI.

The search results panel can be resized by clicking and dragging the lower-right corner of the panel. If the search results exceed the size of the panel, a scroll bar is displayed at the right side of the panel.

The search function is also available within all table views of objects within the GUI. You can filter items displayed in the table by typing a text string in the box located above the table.

The example in Figure 4-35 illustrates the Volume and Snapshots view filtered to display all volumes that contain the “ITSO” string in the name field. The string can be anywhere in the name, not just at the beginning.

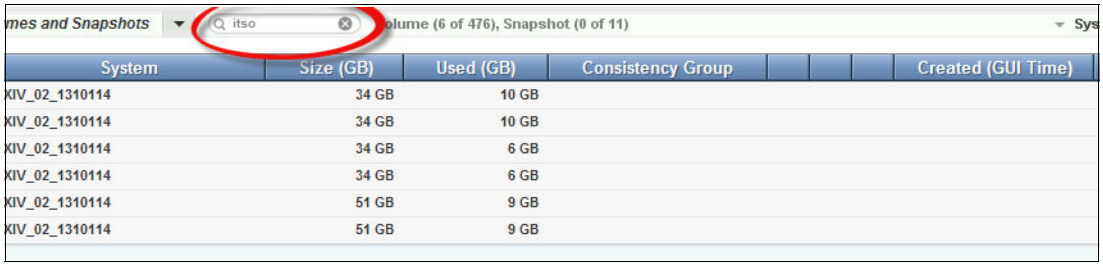


Figure 4-35 Local table filter

Volume identification in hexadecimal

The XIV Storage Management GUI can now be configured to show hexadecimal volume serial numbers rather than decimal volume serial numbers, as shown in Figure 4-36.

To enable this setting, click the **Tools** portion of the menu bar, click **Management**, and select **Hexadecimal** in the Volume Serial drop down field.

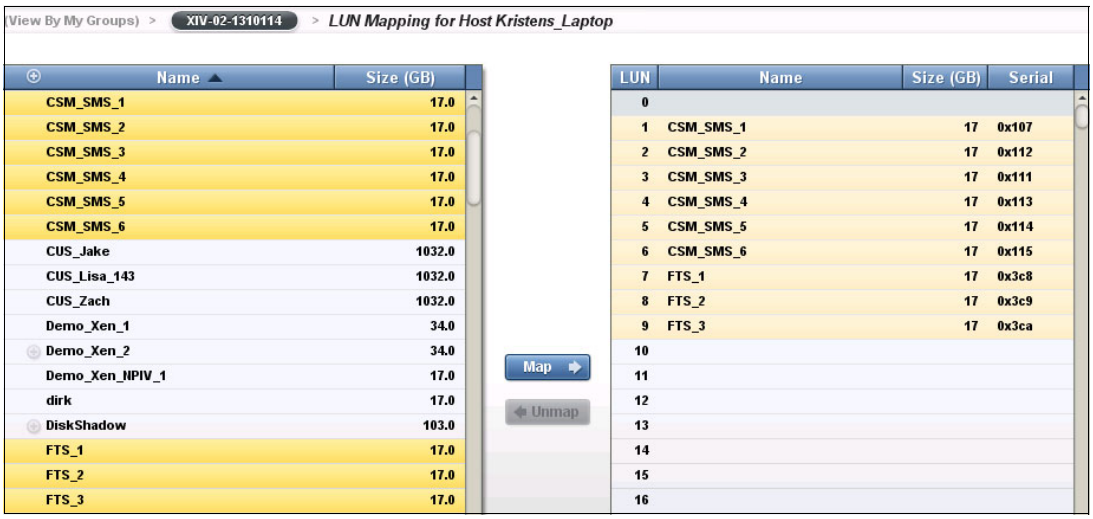


Figure 4-36 Volume identification in hexadecimal

4.3.2 XIV Storage System Command-Line Interface

As with the XIV Storage Management GUI, the XIV Storage System Command-Line Interface (XCLI) is an intuitive tool for managing IBM XIV systems. More details are provided in this section.

Starting the XIV Command-Line Interface

The XCLI Session can be started in Windows by clicking **Start** → **Programs**. Alternatively, if the Create Desktop XCLI icon was selected during the installation, the XCLI Session shortcut is on the desktop. Additionally, an XCLI executable function is available for use from the command line. These various methods of starting the XCLI functions are covered in the following list:

- XCLI session

  Click the XCLI program shortcut in the programs menu or on the desktop. You are prompted for user credentials and the IP address of the XIV system you want to manage. Alternatively, from the XIV GUI, use the XCLI link in the Tools drop-down menu of the menu bar, or right-click in the System view and choose **Launch XCLI** (as shown in Figure 4-37 on page 137). Initiating an XCLI Session from within the XIV Storage Management GUI automatically provides the current user ID and password and connects to the system selected.

  Tip: XCLI Session is the easiest way to issue XCLI commands against XIV Storage Systems.

  Figure 4-37 on page 137 shows the process of starting XCLI from the Systems menu.

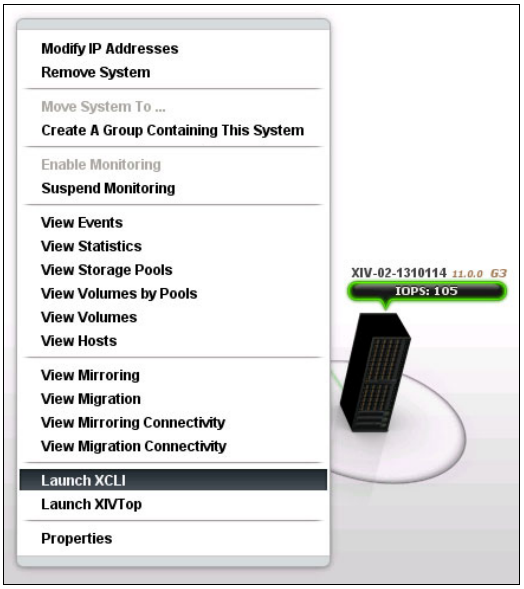


Figure 4-37 Starting XCLI from the Systems menu

- XCLI command prompt

  Another way to start the XCLI is through the Microsoft Windows command prompt. This method can be useful for scripting commands for configuration-related matters. When using this method, user and password information must be provided along with system information, as shown in Figure 4-38.

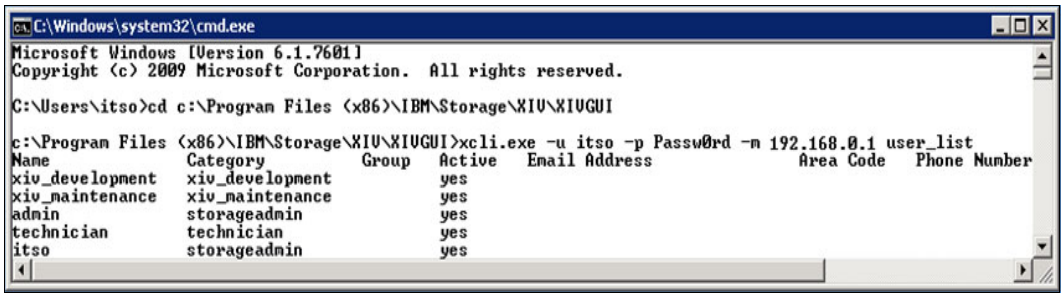

C:\Windows\system32\cmd.exe
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.
C:\Users\itso>cd c:\Program Files (x86)\IBM\Storage\XIV\XIVGUI
c:\Program Files (x86)\IBM\Storage\XIV\XIVGUI>xcli.exe -u itso -p Passw0rd -m 192.168.0.1 user_list
Name Category Group Active Email Address Area Code Phone Number
xiv_development xiv_development yes
xiv_maintenance xiv_maintenance yes
admin storageadmin yes
technician technician yes
itso storageadmin yes

Figure 4-38 Starting XCLI from a Windows command prompt

Tip: For convenience and easier access to XCLI commands, add the following value to the Path system variable:

```
c:\Program Files (x86)\IBM\Storage\XIV\XIVGUI
```

This allows XCLI commands to be run from a command window in any directory.

For more information about using the XCLI, see the *IBM XIV Storage System User Manual*, GC27-3914 and the IBM XIV Storage System documentation in the IBM Knowledge Center:

http://ibm.co/1rvfciG

XCLI session features

XCLI Session offers command and argument completions, along with possible values for the arguments. Within a session, there is no need to enter user information or IP addresses for each command. For example:

- To run a command, type the command.
- For command completion, type part of a command and press Tab to see possible valid commands.
- For command argument completion, type a command and press Tab to see a list of values for the command argument.

Figure 4-39 shows an example of an XCLI Session.

```
>> user_ <TAB>
user_define        user_delete     user_group_add_user    user_group_create
user_group_delete  user_group_list user_group_remove_user user_group_rename
user_group_update  user_list       user_rename            user_update

>> user_list <TAB>
show_users=   user=

>> user_list user= <TAB>
xiv_development   xiv_maintenance   admin    technician
ITSO

>> user_list user=admin
Name   Category     Group Active Email Address Area Code Phone Number Access
All
admin  storageadmin       yes
```

Figure 4-39 XCLI Session example

Customizing the XCLI environment

For convenience and more efficiency in using the XCLI, use the XCLI Session environment and start XCLI Session from the XIV Storage Management GUI menu. However, if you want to write scripts to run XCLI commands, it is possible to customize your management workstation environment.

As part of the XIV Storage System high-availability features, each system is assigned three IP addresses. When running a command, the XCLI utility is provided with these three IP addresses and tries each of them sequentially until communication with one of the IP addresses is successful. You must pass at least one of the IP addresses (IP1, IP2, and IP3) with each command. To avoid redundantly typing and recalling IP addresses, use a predefined configuration name. By default, XCLI uses the system configurations defined when adding systems to the XIV Storage Management GUI. To list the current configurations, use the command shown in Example 4-1.

Example 4-1 List Configurations XCLI command excerpt

```
c:\Users\itso>xcli -L
System             Managements  IPs
XIV LAB 03 1300203 192.168.0.1, 192.168.0.2, 192.168.0.3
XIV 02 1310133     192.168.0.4, 192.168.0.5, 192.168.0.6
```

Commands: When running a command, you must specify either a configuration or IP address, but not both.

To issue a command against a specific XIV Storage System, you must supply the user name and the password for that system. The default user is `admin` and the default password is `adminadmin`, which can be used with the following parameters:

`-u user` or **`-user user`**	Sets the user name that is used to run the command.
`-p password` or **`-password password`**	The XCLI password that must be specified to run a command in the system.
`-m IP1 [-m IP2 [-m IP3]]`	Defines the IP addresses of the XIV Storage System.

Example 4-2 illustrates a common command execution syntax on a given XIV Storage System.

Example 4-2 Simple XCLI command

```
c:\Users\itso>xcli -u itso -p PasswOrd -m 192.168.0.1 user_list
```

Managing the XIV Storage System by using the XCLI always requires that you specify these same parameters. To aid in using them, define and use specific environment variables. Open a command prompt window, in this case in Windows 2008 R2, and set values for specific environment variables, as shown in Example 4-3.

Example 4-3 Script file setup commands

```
setx XIV_XCLIUSER=itso
setx XIV_XCLIPASSWORD=PasswOrd
setx XCLI_CONFIG_FILE=%HOMEDRIVE%%HOMEPATH%\My Documents\xcli\xiv_systems.txt
```

The XCLI requires user and password options. So, if user and passwords are not specified, the default environment variables **`XIV_XCLIUSER`** and **`XIV_XCLIPASSWORD`** are used. Also, the `XCLI_CONFIG_FILE` variable file must be populated before setting the environment variable.

The configuration in this example is stored in a file under the user's home directory. A separate file can be specified by **`-f`** or **`--file`** (applicable to configuration creation, configuration deletion, listing configurations, and command execution). Alternatively, the environment variable **`XCLI_CONFIG_FILE`**, if defined, determines the file's name and path. After running the setup commands, the shortened command syntax works as shown in Example 4-4.

Example 4-4 Short command syntax

```
c:\Users\itso>xcli -a XIV2 -m 192.168.0.4 -m 192.168.0.5 -m 192.168.0.6
c:\Users\itso>xcli -c XIV2 user_list
```

Options: In the previous example, the **`-a`** option is used to name the XIV Storage System designated with the **`-m`** addresses as XIV2, and referred to accordingly afterward.

It is also possible to create set credentials with **xcli --protect**. This command prompts you for the user and password and then stores them in a secure file in the system. The defined default credentials are for the XCLI installation that is used. These credentials are used when no credentials are supplied by using **-u** and **-p** or defined as environment variables. See Example 4-5.

Example 4-5 xcli --protect

```
C:\Users\IBM_ADMIN>xcli --protect
User name: itso
Password:
Confirm password:
Credentials saved locally.
```

Getting help with XCLI commands

To get help about the use of commands (assuming that you have set and logged in as a valid user), proceed as shown in Example 4-6.

Example 4-6 XCLI help commands

```
c:\Users\itso>xcli
c:\Users\itso>xcli -c XIV2 help
c:\Users\itso>xcli -c XIV2 help command=user_list format=full
```

The first command prints the use of XCLI. The second prints all of the commands that can be used in that particular system. The third shows the use of the **user_list** command with all of the parameters.

There are various parameters to get the result of a command in a predefined format. The default is the user readable format. Specify the **-s** parameter to get it in a comma-separated format or specify the **-x** parameter to obtain an XML format.

Fields: The XML format contains all the fields of a particular command. The user and the comma-separated formats provide just the default fields as a result.

To list the field names for a specific **xcli** command, use the **-t** parameter, as shown in Example 4-7.

Example 4-7 XCLI field names

```
c:\Users\itso>xcli -c XIV2 -t name,fields help command=user_list
```

Scripts

XIV Storage Management software XCLI commands can be used in scripts or batch programs if you need to use repetitive or complex operations. The XCLI can be used either in a shell environment to interactively configure the system or as part of a script to perform specific tasks, as shown in Example 4-3 on page 139. In general, the XIV Storage Management GUI or the XCLI Session environment nearly eliminate the need for scripts.

4.4 Storage pools

We introduced the concept of XIV storage pools in 2.6, “Storage pool concepts” on page 35.

4.4.1 Function of storage pools

Storage pools manage a related group of logical volumes and their snapshots. Storage pools offer the following key benefits:

- Improved management of storage space

  Specific volumes can be grouped within a storage pool to give you the flexibility to control the use of storage space by specific applications, a group of applications, or departments.

- Improved regulation of storage space

  Automatic snapshot deletion occurs when the storage capacity limit is reached for each storage pool. Therefore, when a storage pool’s size is exhausted, only the snapshots that are in the affected storage pool are deleted.

Size of storage pools

The size of storage pools and the associations between volumes and storage pools are constrained by the following circumstances:

- The size of a storage pool can range from 17 GB up to 325 TB on XIV Storage System Gen3 systems equipped with 4 TB disks.

  Tip: To determine the maximum pool size, use the **`system_capacity_list`** command. Look at the column titled `Max_Pool_Size`.

- The size of a storage pool can always be decreased, limited only by the space already consumed by the volumes and snapshots in that storage pool.
- Volumes can be moved between storage pools without any limitations, if they are not part of a consistency group and there is enough free space in the target storage pool. Volumes that are part of a consistency group can be moved together as a group.

  Important: All of these operations are handled by the system at the metadata level, and they do not cause any data movement (copying) from one disk drive to another. They are completed almost instantly and can be done at any time without affecting the applications.

Thin-provisioned pools

Thin provisioning is the practice of allocating storage on a “just-in-time” and “as needed” basis by defining a logical, or soft, capacity that is larger than the physical, or hard, capacity. Thin provisioning enables XIV Storage System administrators to manage capacity based on the total space consumed rather than just the space allocated.

Thin provisioning can be specified at the storage pool level. Each thin-provisioned pool has its own hard capacity (which limits the actual disk space that can be consumed) and soft capacity (which limits the total logical size of volumes defined). The difference in the pool size depends on the type of pool:

- *Hard pool size*: The hard pool size represents the physical storage capacity allocated to volumes and snapshots in the storage pool. The hard size of the storage pool limits the total of the hard volume sizes of all volumes in the storage pool and the total of all storage consumed by snapshots.
- *Soft pool size*: This size is the limit on the total soft sizes of all the volumes in the storage pool. The soft pool size has no effect on snapshots.

For more information about the concept of thin provisioning and a description of hard and soft size for storage pools and volumes, see 2.7, “Capacity allocation and thin provisioning” on page 37, and the Redpaper publication titled *IBM XIV Storage System Thin Provisioning and Space Reclamation*, REDP-5001.

When using the XIV Storage Management GUI, you specify what type of pool is wanted (Regular Pool or a Thin Pool) when creating the pool. See “Creating storage pools” on page 144. When using the XCLI, you create a thinly provisioned pool by setting the soft size to a value greater than its hard size.

If the requirements for the pool change later on, the pool’s type can be changed (nondisruptively).

Tip: Thin provisioning management is performed individually for each storage pool, and running out of space in one pool does not affect other pools.

4.4.2 Managing storage pools with the XIV Storage Management GUI

In this section, we describe how to manage storage pools using the XIV Storage Management GUI.

Important: Illustrations in this chapter mostly apply to an XIV Storage System fully configured with 1 TB drives.

Managing pools with the XIV Storage Management GUI is fairly simple and intuitive. The related tasks can be reached either through the menu bar or the corresponding function icon on the left (called *Pools*), as shown in Figure 4-40.

Figure 4-40 Opening the Pools menu

To view overall information about the storage pools, click **Storage Pools** from the Pools menu that is shown in Figure 4-40. This opens the Storage Pools view, as shown in Figure 4-41 on page 143.

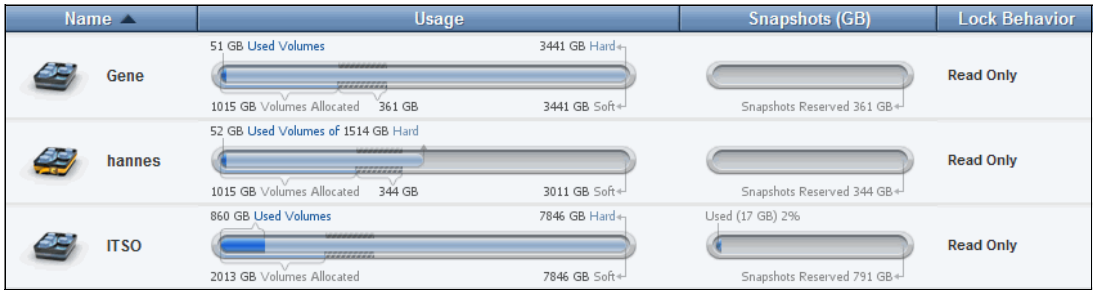


Figure 4-41 Storage Pools view

The Storage Pools view displays a table of all the pools in the system, with a series of gauges for each pool. This view gives the administrator a quick grasp and general overview of essential information about the system pools. The Storage Pools view can be customized, allowing you to show or hide specific columns. The Name and Usage columns are required. The following optional column values are available:

- Creator
- Hard (Free)
- Hard (Total)
- Lock Behavior
- Snapshots (GB)
- Snapshots (Total)
- Snapshots Used
- Soft (Free)
- Soft (Total)
- System
- Volumes Used

The capacity consumption by volumes and snapshots within a particular storage pool is indicated by various colors. The default threshold values are as follows:

- Blue indicates consumed capacity below 80%.
- Yellow indicates capacity consumption above 80%.
- Orange indicates capacity consumption of over 90%.
- Red indicates that a storage pool has depleted hard capacity.

The name, the size, and the separated segments are labeled appropriately.

In Figure 4-42 on page 144, you can see an example of storage pool figures. The figure depicts the following values for a storage pool:

Used volumes Physical amount of data already written to the storage pool. It is also the sum of space consumed on each of the volumes in the pool.

Volumes allocated Logical amount of space reserved for all defined volumes in the storage pool.

Hard pool size Storage pool hard limit. It represents the physical storage capacity allocated to volumes and snapshots in the storage pool. See 4.4.1, “Function of storage pools” on page 141 for a description.

Soft pool size Storage pool soft limit. It is the limit for the total soft sizes of all the volumes in the storage pool. See 4.4.1, “Function of storage pools” on page 141 for a description.

Figure 4-42 on page 144 provides an example of storage pool and size numbers.

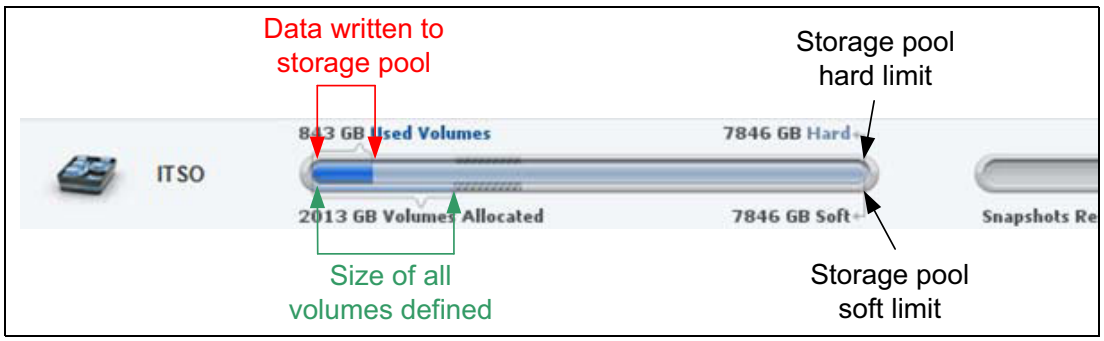


Figure 4-42 Storage pool and size numbers

To export the details shown in the Storage Pools view, click the **Export** icon in the toolbar at the top of the GUI window. For more information, see “Copy and paste of system configuration settings” on page 129.

Creating storage pools

The creation and resizing of storage pools is straightforward, and you need to only take care with the size allocation and reallocation. The name of a storage pool must be unique in the system.

Pool size: The size of the storage pool is specified as an integer multiple of 10^9 bytes, but the actual size of the created storage pool is rounded up to the nearest integer multiple of $16x2^{30}$ bytes. According to this rule, the smallest pool size is 17 GB.

When creating a storage pool, the system initially provides a default snapshot size of 10%. However, you might want to adjust this value, depending on how snapshots are used, your workload characteristics, and your requirements for how long snapshots must be kept. This value can be set at the time of creation or dynamically changed later, depending on your needs.

Snapshot size: The snapshot size (default or specified) is a subset of the specified pool size. It does not allocate more space.

Sizing must consider volumes that are to be added to (or exist in) the specific storage pool, the current allocation of storage in the total system capacity, and future activity within the storage pool, especially regarding snapshot propagation from creating too many snapshots. Upon depletion of space in a pool, the system progressively deletes snapshots in the pool to free up space for additional write requests. For more information, see Chapter 1, “Snapshots,” in *IBM XIV Storage System Business Continuity Functions*, SG24-7759.

The system enables the assignment of the entire available capacity to user-created storage pools. The storage pool is initially empty and does not contain volumes. However, you cannot create a storage pool with zero capacity.

To create a storage pool, complete the following steps:

1. Click **Add Pool** in the toolbar from in the Storage Pools view, or right-click in an empty area of the Storage Pools view and click **Add Pool**. An Add Pool window opens, as shown in Figure 4-43.

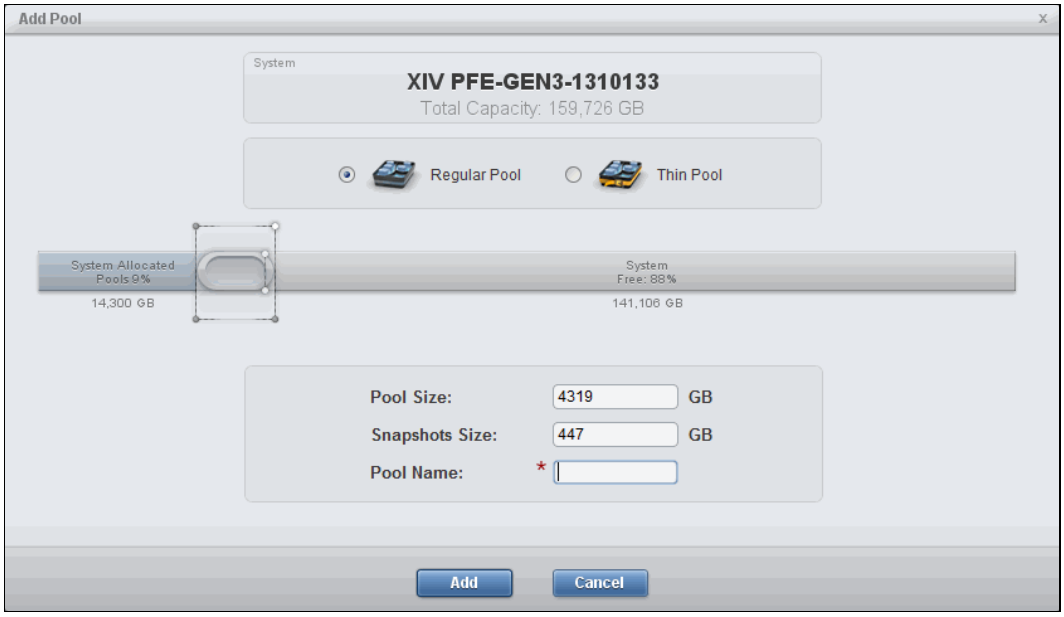


Figure 4-43 Add Pool

2. You must choose **Regular Pool** or **Thin Pool** according to your needs. Based on the pool type you choose, the available fields differ. For a Thin Pool, enter values for the following fields:
   - Pool Hard Size: Specify the upper limit of hard capacity.
   - Pool Soft Size: Specify the upper limit of soft capacity.
   - Lock Behavior: Specify the behavior in case of depleted capacity.

     This value specifies whether the storage pool is locked for write or whether it is disabled for both read and write when running out of storage space. The default value is *Read only.*

   For a regular pool, enter the Pool Size field, the required size of the storage pool.
3. In the Snapshots Size field, enter the required size of the reserved snapshot area.
4. In the Pool Name field, enter a name (it must be unique across the storage system) for the storage pool.
5. Click **Add** to add this storage pool.

Resizing storage pools

This action can be used to both increase or decrease a storage pool size. Capacity calculation is performed in respect to the total system net capacity. All reductions and increases are reflected in the remaining free storage capacity.

Storage pool size:

- When increasing a storage pool size, you must ensure that the total system capacity holds enough free space for the increase in storage pool size.
- When decreasing a storage pool size, you must ensure that the storage pool has enough free capacity for a reduction in size.

This operation is also used to shrink or increase the snapshot capacity inside the storage pool. This alteration affects only the space within the storage pool. In other words, increasing snapshot size consumes the free capacity only from the corresponding pool.

To change the size of one storage pool in the system, right-click a pool in the Storage Pools view (Figure 4-41 on page 143) and select **Resize**.

The window shown in Figure 4-44 opens. Change the pool hard size, soft size, or the snapshot size to match your new requirements. Within the storage pool gauge, the vertical dotted line to the left (with the blue triangles) is the consumed capacity and the vertical dotted line to the right (with the white circles) is the snapshot size.

Figure 4-44 Resizing pool

The resize operation can also be used to change the type of storage pool from thin-provisioned to regular or from regular to thin-provisioned (see Figure 4-45). Change the type of pool in the Resize Pool window selection:

- When a regular pool is converted to a thin-provisioned pool, you must specify an additional soft pool size parameter besides the existing hard pool size. The soft pool size must be greater than or equal to the hard pool size.
- When a thin-provisioned pool is changed to a regular pool, the soft pool size parameter disappears from the window; in fact, its value is equal to the hard pool size.

  If the space consumed by existing volumes exceeds the pool's actual hard size, the pool cannot be changed to a regular type pool. In this case, you must specify a minimum hard pool size equal to the total capacity consumed by all the volumes within this pool.

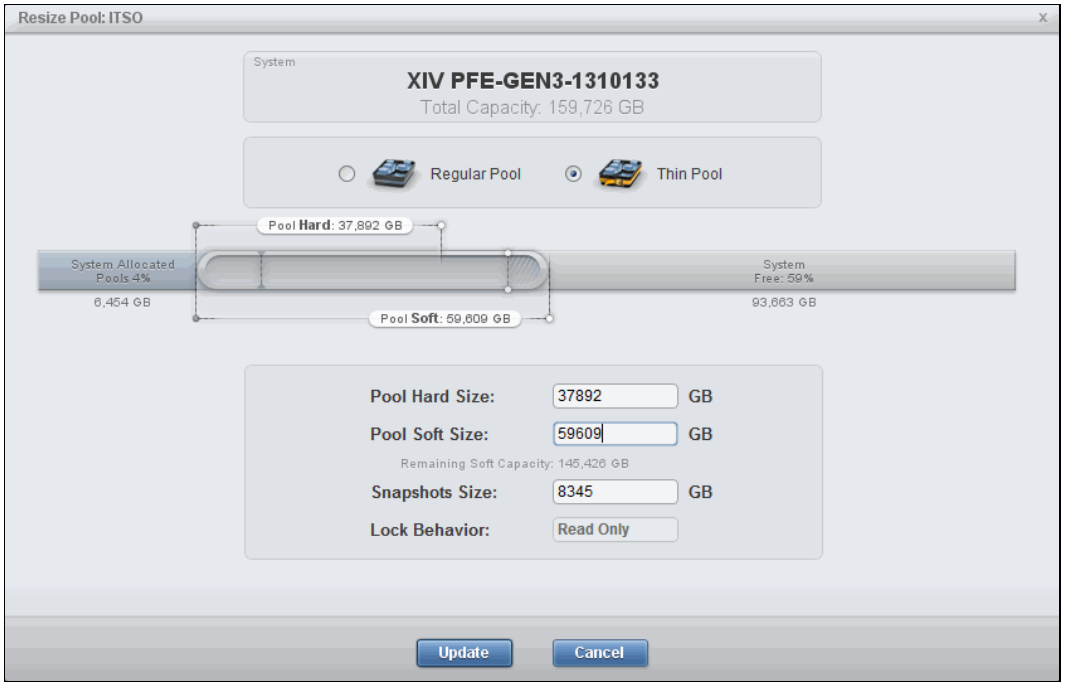


Figure 4-45 Resizing and changing the type of a pool

The remaining soft capacity is displayed under the **Pool Soft Size** setting and calculated by the system in the following manner:

Remaining Soft Capacity = [Current Storage Pool Soft Size + Remaining System Soft Size] - Current Storage Pool Hard Size

Deleting storage pools

To delete a storage pool, right-click the storage pool and select **Delete**. The system asks for a confirmation before deleting this storage pool.

The capacity of the deleted storage pool is reassigned to the system's free capacity, which means that the free hard capacity increases by the size of the deleted storage pool.

Consideration: You cannot delete a storage pool if it still contains volumes.

Moving volumes between storage pools

In order for a volume to be moved to a specific storage pool, there must be enough room for the volume to be there. If there is not enough free capacity (meaning that adequate capacity has not been allocated), the storage pool must be resized, or other volumes must be moved out first to make room for the new volume.

When moving a master volume from one storage pool to another, all of its snapshots are moved along with it to the destination storage pool. You cannot move a snapshot alone (independent of its master volume).

The destination storage pool must have enough free storage capacity to accommodate the volume and its snapshots. The exact amount of storage capacity allocated from the destination storage pool is released at the source storage pool.

A volume that belongs to a consistency group cannot be moved without moving the entire consistency group.

As shown in Figure 4-46, in the Volume by Pools view, right-click the appropriate volume and initiate a **Move to Pool** operation to change the location of a volume.

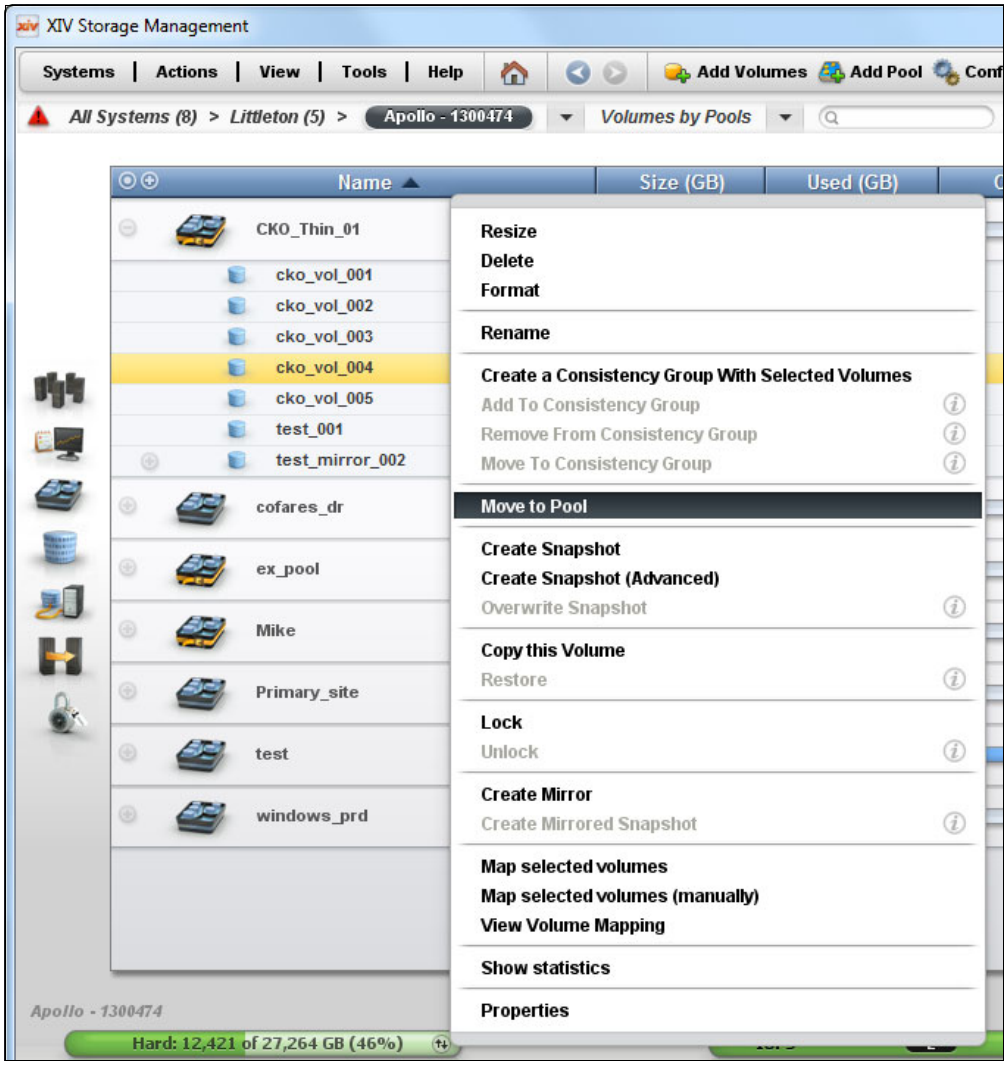


Figure 4-46 Volumes by Pools view

In the pop-up window, select the appropriate storage pool, as shown in Figure 4-47. Click **OK** to move the volume into this pool.

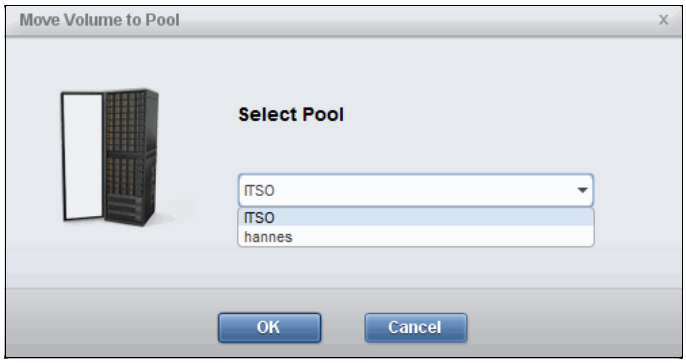


Figure 4-47 Move volume to another pool

Pool alert thresholds

You can use the XIV Storage Management GUI to configure pool use thresholds to trigger color-coded visual alerts at various severity levels. These alerts are either system-based or, starting with XIV Version 11.5 and GUI version 4.4, thresholds can be set at a pool level (and, therefore, controlled by domains). The defined thresholds are viewable in the XIV Storage Management GUI by all users who are defined for the particular system. System events are also created based on the specified values. The default pool use threshold values are for warning (80%), minor (90%), and major (95%).

To configure pool thresholds from the XIV Storage Management GUI Storage Pools view, right-click an empty area in the view and choose **Pool Thresholds** to open a dialog window, as shown in Figure 4-48.

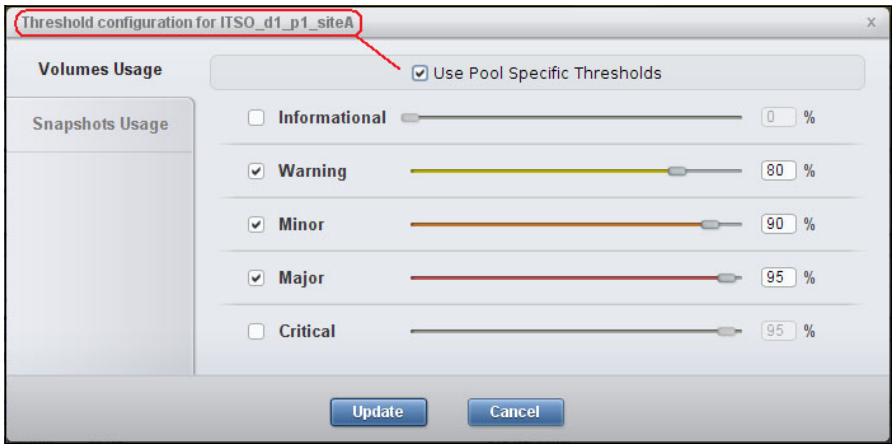


Figure 4-48 Set pool alert thresholds

Enable a specific class of Volumes Usage or Snapshots Usage threshold alert by checking the box next to it. To adjust the value of an enabled threshold, click and drag the slider bar or type in the corresponding numerical value.

Important: When pool-specific thresholds are enabled, they override or take precedence on the system level thresholds.

Events: The XIV Storage System has several predefined events that are triggered when system pool use is depleting. Rules can be set to receive email alerts for these events. For more information about setting up email event notifications, see Chapter 7, “Monitoring” on page 327.

The system wide pool thresholds can be copied from one XIV system to another by copying the system configuration of the source system and then selecting Paste Pool Threshold from the target system pop-up menu. The specific pool thresholds can be copied from one pool to another within the same XIV system, but cannot be copied from one XIV system to another XIV system.

4.4.3 Managing storage pools with XIV Command-Line Interface

All of the operations described in 4.4.2, “Managing storage pools with the XIV Storage Management GUI” on page 142 can also be accomplished through the XIV command-line interface (XCLI).

Listing pool-related commands

To get a list of all the storage pool-related XCLI commands, enter the following command from the XCLI command shell:

```
help category=storage-pool
```

Important: The commands shown in this section are based on the assumption that you started an XCLI session on the selected system, as described in “XCLI session features” on page 138.

The output that is shown in Example 4-8 is displayed.

Example 4-8 All the storage pool-related commands

```
Category      Name                    Description
storage-pool  cg_move                 Moves a Consistency Group, all its volumes and all their snapshots
                                          and Snapshot Sets from one Storage Pool to another.
storage-pool  pool_change_config      Changes the lock behavior of the storage Pool.
storage-pool  pool_config_snapshots   Changes the Storage Pool snapshot settings.
storage-pool  pool_create             Creates a Storage Pool.
storage-pool  pool_delete             Deletes a Storage Pool.
storage-pool  pool_list               Lists all Storage Pools or the specified one.
storage-pool  pool_rename             Renames a specified Storage Pool.
storage-pool  pool_resize             Resizes a Storage Pool.
storage-pool  vol_move                Moves a volume and all its snapshot from one Storage Pool to
                                          another.
```

To list the existing storage pools in a system, run **`pool_list`**. A sample output of this command is shown in Figure 4-49.

```
.                  Soft      Snap      Soft       Hard      Hard
Name       Size(GB) Vols(GB)  Size(GB)  Empty(GB)  Size(GB)  Vols(GB)  Locked
Pete       3011     1015      344       1651       1514      51        no
ITSO       7846     2288      791       4766       7846      946       no
Gene       3441     1256      361       1824       3441      206       no
SLES-ITSO  206      103       34        68         206       0         no
```

Figure 4-49 Result of the pool_list command

To create a pool, enter the following command:

```
pool_create pool=”ITSO Pool 2” size=515 snapshot_size=103
```

The size of the storage pool is specified as an integer multiple of 10^9 bytes, but the actual size of the created storage pool is rounded up to the nearest integer multiple of $16x2^{30}$ bytes. The **snapshot_size** parameter specifies the size of the snapshot area within the pool. It is a mandatory parameter, and you must specify a positive integer value for it.

Run the following command to resize an existing pool:

```
pool_resize pool="ITSO Pool 2" size=704 snapshot_size=154
```

With this command, you can increase or decrease the pool size. The **pool_create** and the **pool_resize** commands are also used to manage the size of the snapshot area within a storage pool.

To rename an existing pool, run the following command:

```
pool_rename pool="ITSO Pool 2" new_name="ITSO Pool"
```

To delete a pool, run the following command:

```
pool_delete pool="ITSO Pool"
```

Approve or deny deletion by responding `y/n` when prompted or use the **-y** parameter with the **pool_delete** command to approve deletion.

Tip: You can use the **-y** parameter at the end of a command that requires confirmation to have it auto-approved. This parameter is useful for scripting.

Run the following command to move the volume named `Zejn_02` to `ITSO Pool 3`:

```
vol_move pool="ITSO Pool 3" vol="Zejn_02"
```

The command succeeds only if the destination storage pool has enough free storage capacity to accommodate the volume and its snapshots. The command moves a particular volume and its snapshots from one storage pool to another one. However, if the volume is part of a consistency group, the entire group must be moved. In this case, the **cg_move** command is the correct solution:

```
cg_move cg="Mainz01 CG" pool="ITSO Pool 3"
```

All volumes, volume snapshots, and snapshot groups of the consistency group are moved.

Thinly provisioned pools

To create thinly provisioned pools, specify the **hard_size** and the **soft_size** parameters. For thin provisioning concepts, see 2.7, “Capacity allocation and thin provisioning” on page 37.

A typical storage pool creation command with thin provisioning parameters can be issued as shown in the following example:

```
pool_create pool="ITSO Pool" hard_size=807 soft_size=1013 lock_behavior=read_only
snapshot_size=206
```

The **soft_size** parameter is the maximal storage capacity seen by the host and cannot be smaller than the **hard_size** parameter, which is the hard physical capacity of the storage pool. If a storage pool runs out of hard capacity, all of its volumes are locked to all write commands. Even though write commands that overwrite existing data can be technically serviced, they are blocked to ensure consistency.

To specify the **lock_behavior** parameter, which is the behavior when capacity reserve has been depleted in a thin-provisioned pool, run the following command:

```
pool_change_config pool="ITSO Pool" lock_behavior=no_io
```

This command specifies whether the storage pool is locked for write or whether it disables both read and write when running out of storage space.

Lock: The **lock_behavior** parameter can be specified for non-thin provisioning pools, but it has no effect.

4.5 Volumes

After defining storage pools, the next logical step in system configuration is volume management.

The XIV Storage System offers logical volumes as the basic data storage element for allocating usable storage space to attached hosts. This logical unit concept is known and is widely used by other storage subsystems and vendors. However, the volume segmentation and its distribution over the physical disks is not conventional in the XIV Storage System.

Traditionally, logical volumes are defined within various Redundant Array of Independent Disks (RAID) arrays, where their segmentation and distribution are manually specified. The result is often a suboptimal distribution within and across modules (expansion units) and is dependent upon the administrator's knowledge and expertise.

As explained in 2.3, "Full storage virtualization" on page 24, the XIV Storage System uses true virtualization as one of the basic principles for its unique design. With the XIV Storage System, each volume is divided into 1 MB partitions, and these partitions are distributed pseudo randomly and evenly, and duplicated for protection. The result is optimal distribution in and across all modules, which means that for any volume, the physical drive location and data placement are invisible to the user. This method dramatically simplifies storage provisioning, letting the system automatically lay out the user's volume in an optimal way.

This method offers complete virtualization, without requiring preliminary volume layout planning or detailed and accurate stripe or block size pre-calculation by the administrator. All disks are equally used to maximize the I/O performance and use all the processing power and all the bandwidth available in the storage system.

XIV Storage System virtualization incorporates an advanced snapshot mechanism with unique capabilities, which enables the creation of a virtually unlimited number of point-in-time copies of any volume, without incurring any performance penalties. The concept of snapshots is described in detail in the *IBM XIV Storage System Product Overview*, GA32-0791, and *IBM XIV Storage System Business Continuity Functions*, SG24-7759. You can download the *Product Overview* from the IBM Knowledge Center:

http://ibm.co/1s999AG

Volumes can also be grouped into larger sets called *consistency groups* and *storage pools*. See 2.6, "Storage pool concepts" on page 35.

Important: As shown in Figure 4-50, the basic hierarchy of the logical storage concepts of an XIV Storage System is as follows:

- A volume can have multiple snapshots.
- A volume can be part of one and only one consistency group.
- A volume and its corresponding snapshots are always part of only one storage pool.
- All volumes of a consistency group must belong to the same storage pool.

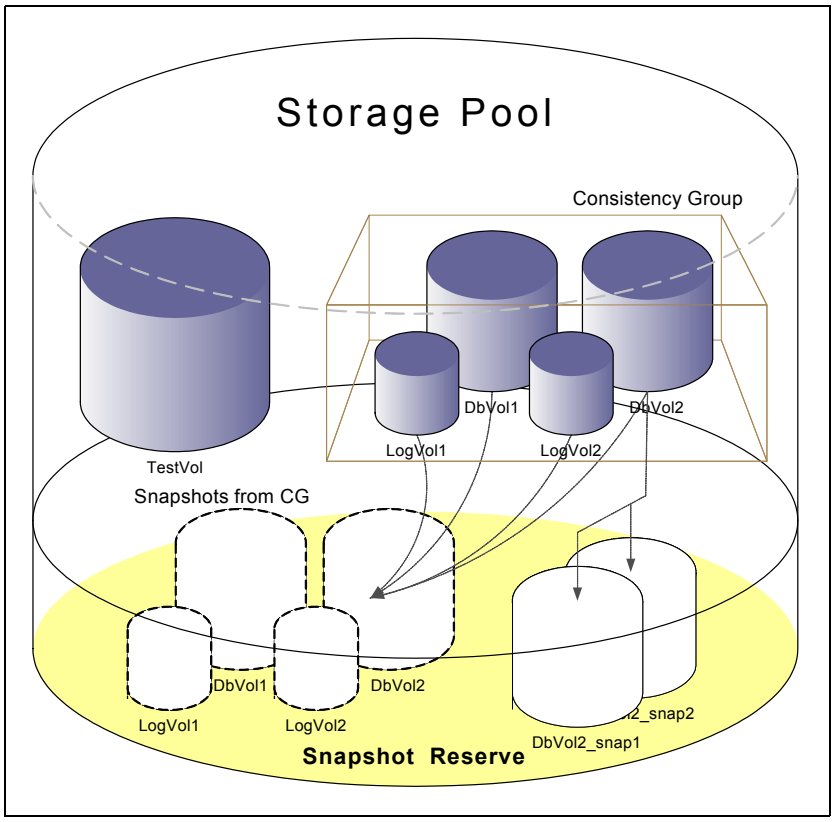


Figure 4-50 Basic storage hierarchy

4.5.1 Managing volumes with the XIV Storage Management GUI

To start a volume management function from the XIV Storage Management GUI, you can either click **View** → **Volumes** → **Volumes and Snapshots** from the menu bar or mouse over the **Volumes** icon and then select the appropriate menu item. See Figure 4-51.

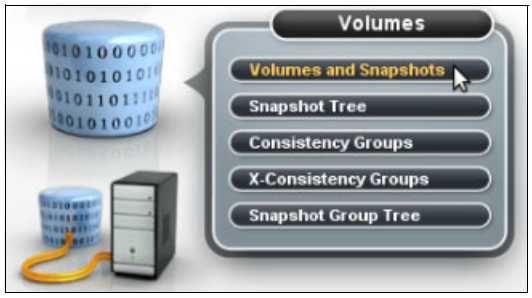
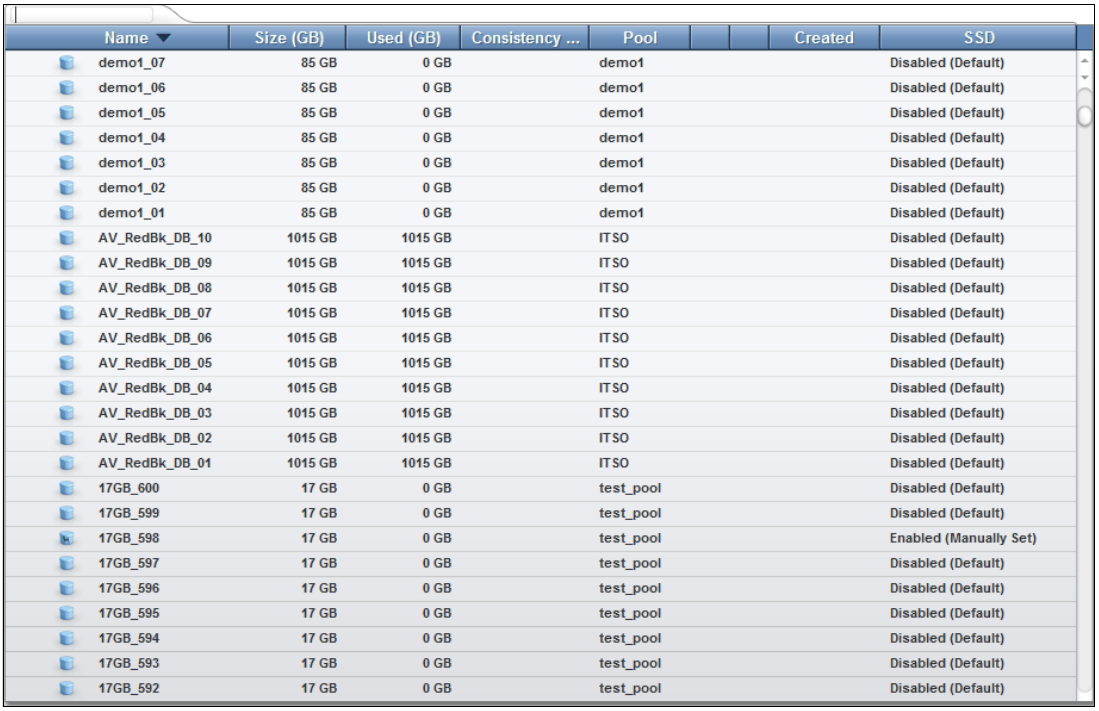


Figure 4-51 Opening the Volumes menu

The Volumes and Snapshots menu item is used to list all the volumes and snapshots that have been defined in this particular XIV Storage System. An example of the resulting window is shown in Figure 4-52.



Name	Size (GB)	Used (GB)	Consistency ...	Pool	Created	SSD
demo1_07	85 GB	0 GB		demo1		Disabled (Default)
demo1_06	85 GB	0 GB		demo1		Disabled (Default)
demo1_05	85 GB	0 GB		demo1		Disabled (Default)
demo1_04	85 GB	0 GB		demo1		Disabled (Default)
demo1_03	85 GB	0 GB		demo1		Disabled (Default)
demo1_02	85 GB	0 GB		demo1		Disabled (Default)
demo1_01	85 GB	0 GB		demo1		Disabled (Default)
AV_RedBk_DB_10	1015 GB	1015 GB		ITSO		Disabled (Default)
AV_RedBk_DB_09	1015 GB	1015 GB		ITSO		Disabled (Default)
AV_RedBk_DB_08	1015 GB	1015 GB		ITSO		Disabled (Default)
AV_RedBk_DB_07	1015 GB	1015 GB		ITSO		Disabled (Default)
AV_RedBk_DB_06	1015 GB	1015 GB		ITSO		Disabled (Default)
AV_RedBk_DB_05	1015 GB	1015 GB		ITSO		Disabled (Default)
AV_RedBk_DB_04	1015 GB	1015 GB		ITSO		Disabled (Default)
AV_RedBk_DB_03	1015 GB	1015 GB		ITSO		Disabled (Default)
AV_RedBk_DB_02	1015 GB	1015 GB		ITSO		Disabled (Default)
AV_RedBk_DB_01	1015 GB	1015 GB		ITSO		Disabled (Default)
17GB_600	17 GB	0 GB		test_pool		Disabled (Default)
17GB_599	17 GB	0 GB		test_pool		Disabled (Default)
17GB_598	17 GB	0 GB		test_pool		Enabled (Manually Set)
17GB_597	17 GB	0 GB		test_pool		Disabled (Default)
17GB_596	17 GB	0 GB		test_pool		Disabled (Default)
17GB_595	17 GB	0 GB		test_pool		Disabled (Default)
17GB_594	17 GB	0 GB		test_pool		Disabled (Default)
17GB_593	17 GB	0 GB		test_pool		Disabled (Default)
17GB_592	17 GB	0 GB		test_pool		Disabled (Default)

Figure 4-52 Volumes and Snapshots view

Volumes are listed in a tabular format. If the volume has snapshots, then a + or a - icon displays on the left. Snapshots are listed under their master volumes, and the list can be expanded or collapsed at the volume level by clicking the + or - icon.

Snapshots are listed as a subbranch of the volume of which they are a replica, and their row is indented and slightly shaded.

The Master column of a snapshot shows the name of the volume of which it is a replica. If this column is empty, the volume is the master.

Tip: To customize the columns in the lists, right-click one of the column headings and make the required selection of attributes. The default column set does not contain the Master column. You can also resize the columns to allow for longer names or to make more columns visible.

Table 4-1lists the columns in the Volumes and Snapshots view.

Table 4-1 Columns in the Volumes and Snapshots view

Column	Description	Optional
Name	Name of a volume or snapshot.	N
System	Name of the XIV Storage System to which this volume or snapshot belongs.	Y
Size (GB)	Volume or snapshot size. (The value is zero if the volume is specified in blocks.)	Y

Column	Description	Optional
Used (GB)	Used capacity in a volume	Y
Size (Blocks)	Volume size in blocks	Y
Size (Disk)	Volume size in GB	Y
Used Capacity (MiB)	Consumed capacity in binary mebibytes (2^{20} bytes)	Y
Master	Snapshot master's name	Y
Consistency Group	Consistency group name	Y
Pool	Storage pool name	Y
SSD (solid-state drive)	Shows SSD caching status	Y
Locked Status	Indicates the locking status of a volume or snapshot	Y
Snapshot Modification Status	Shows if the snapshot was unlocked or modified	Y
Coupling Status	Shows the copy services coupling status	Y
Snapshot Formatted	Shows if the snapshot has been formatted	Y
Deletion Priority	Indicates the priority of deletion by numbers for snapshots	Y
Created (GUI Time)	Shows the creation time of a snapshot as measured at the GUI	Y
Created (System Time)	Shows the creation time of a snapshot as measured at the XIV Storage System	Y
Created on Master (GUI Time)		Y
Created on Master (System Time)		Y
Creator	Volume or snapshot creator name	Y
Serial Number	Volume or snapshot serial number	Y

Most of the volume-related and snapshot-related actions can be selected by right-clicking any row in the table to display a drop-down menu of options. The options in the menu differ slightly for volumes and snapshots.

Menu option actions

The following actions can be performed through these menu options:

- Adding or creating volumes (see "Creating volumes" on page 156)
- Resizing a volume (see "Resizing volumes" on page 160)
- Deleting a volume or snapshot (see "Deleting volumes" on page 161)
- Formatting a volume
- Renaming a volume or snapshot
- Creating a consistency group with these volumes
- Adding to a consistency group
- Removing from a consistency group

- Moving volumes between storage pools
  (See “Moving volumes between storage pools” on page 148)
- Creating a snapshot
- Creating a snapshot (advanced)
- Overwriting a snapshot
- Copying a volume or snapshot
- Locking and unlocking a volume or snapshot
- Mappings
- Change SSD caching state
- Displaying properties of a volume or snapshot
- Changing a snapshot’s deletion priority
- Duplicating a snapshot or a snapshot (advanced)
- Restoring from a snapshot

Creating volumes

When you create a volume in a *traditional* or *regular* storage pool, the entire volume storage capacity is reserved (static allocation). You cannot define more space for volumes in a regular storage pool than the actual hard capacity of the pool, which guarantees the functions and integrity of the volume.

When you create a volume in a *thin-provisioned* pool, the capacity of the volume is not reserved immediately to the volumes. However, a basic 17.1 GB piece, which is taken out of the storage pool hard capacity, is allocated at the first I/O operation. In a thin-provisioned pool, you are able to define more space for volumes than the actual hard capacity of the pool, up to the soft size of the pool.

The volume size is the actual “net” storage space, as seen by the host applications, not including any mirroring or other data protection impact. The free space consumed by the volume is the smallest multiple of 17 GB that is greater than the specified size. For example, if we request an 18 GB volume to be created, the system rounds this volume size to 34 GB. For a 16 GB volume size request, it is rounded to 17 GB.

Figure 4-53 on page 157 gives you various basic examples of volume definition and planning in a thinly provisioned pool.

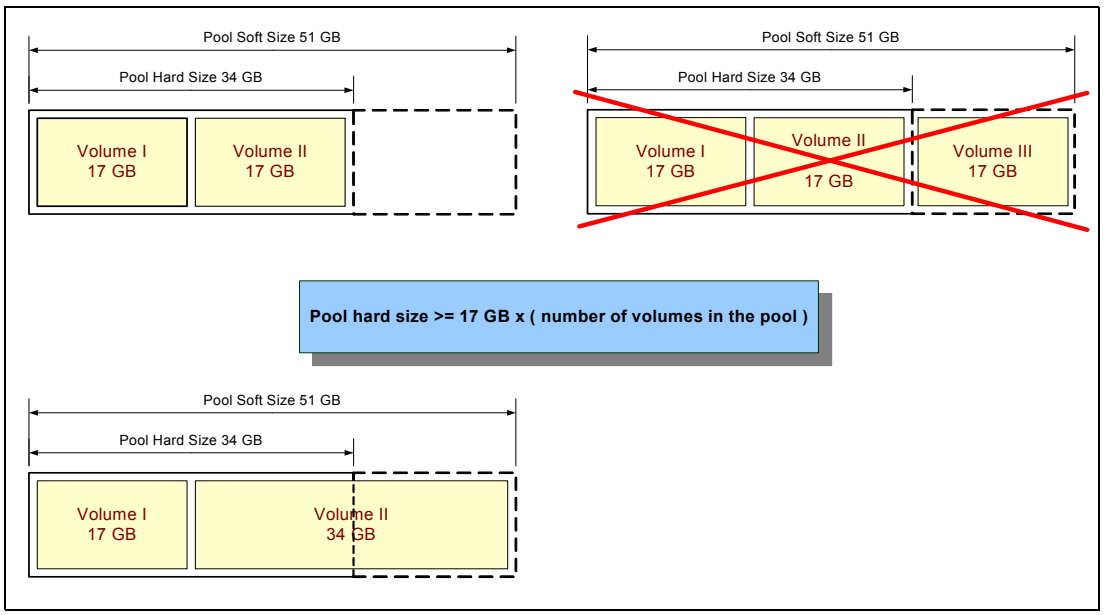


Figure 4-53 Planning the number of volumes in a thin-provisioned pool

Figure 4-53 shows the volumes with the minimum amount of capacity, but the principle can be used for larger volumes as well. Plan carefully the number of volumes or the hard size of the thinly provisioned pool because of the minimum hard capacity that is consumed by one volume.

If you create more volumes in a thinly provisioned pool than the hard capacity can cover, the I/O operations against the volumes fail at the first I/O attempt.

Volumes: Plan the volumes in a thin-provisioned pool in accordance with this formula:

```
Pool Hard Size >= 17 GB x (number of volumes in the pool)
```

The size of a volume can be specified either in gigabytes (GB), gibibytes (GiB), or in blocks (where each block is 512 bytes). If the size is specified in blocks, volumes are created in the exact size specified, and the size is not rounded up. It means that the volume shows the exact block size and capacity to the hosts but consumes a 17 GB size in the XIV Storage System. This capability is relevant and useful in migration scenarios.

If the size is specified in GB, the actual volume size is rounded up to the nearest 17.1 GB multiple (making the actual size identical to the free space consumed by the volume). This rounding up prevents a situation where storage space is not fully used because of a gap between the free space used and the space available to the application.

If the size is specified in GiB, the volume is specified in binary gigabytes (in multiples of 2^{30}) and is rounded up to the nearest 16 GiB multiple, which is physically the same total size as the equivalent reported gigabyte size.

GB versus GiB (and XIV Storage System volumes):

- A gigabyte is a decimal multiple of a byte unit. 1 GB equals 10^9 bytes (1,000,000,000 bytes).
- A gibibyte is a binary multiple of a byte unit. 1 GiB equals 2^{30} bytes (or 1,073,741,824 bytes).

When you create a volume on an XIV Storage System, it allocates multiples of 17.1 GB/16 GiB.

The volume is logically formatted at creation time, which means that any read operation results in returning all zeros as a response.

To create volumes with the XIV Storage Management GUI, complete the following steps:

1. Click the **Add Volumes** icon in the Volume and Snapshots view (see Figure 4-52 on page 154) or right-click in the body of the window (not on a volume or snapshot) and click **Add Volumes**.

   The window that is shown in Figure 4-54 opens.

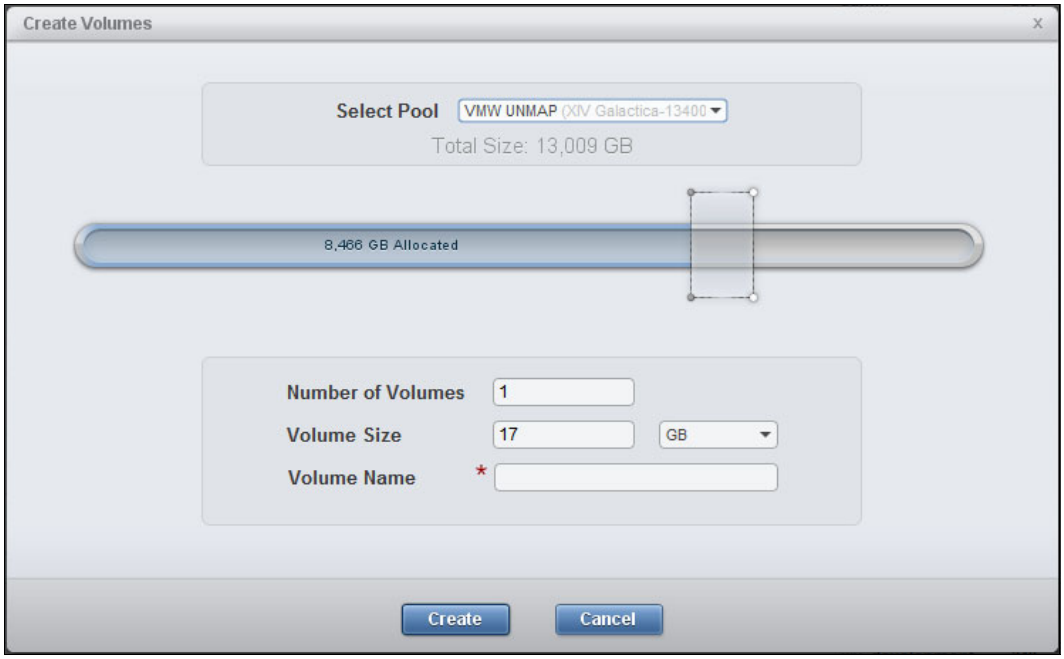


Figure 4-54 Create Volumes view

2. From the Select Pool field, select the pool where this volume should be stored. See 4.4.2, “Managing storage pools with the XIV Storage Management GUI” on page 142 for a description of how to define storage pools. The storage size and allocation of the selected storage pool is shown textually and graphically:
   - The blue portion of the bar indicates the space already allocated in this storage pool.
   - The shaded portion of the bar, outlined with a rectangular size indicator, indicates the space that will be allocated to this volume (or volumes) after it is created.
   - The remaining gray portion of the bar indicates the space that remains free after this volume (or volumes) is allocated.
3. In the Number of Volumes field, specify the required number of volumes.

4. In the Volume Size field, specify the size of each volume to define. The size can also be modified by dragging the right side of the rectangular size indicator.

   Volume size: When multiple volumes are created in the same step, they all have the same size as specified in the Volume Size field.

5. In the Volume Name field, specify the name of the volume to define. The name of the volume must be unique in the system. If you specified that more than one volume is defined, they are automatically and successively named by appending an incrementing number to end of the specified name. You can manually input the suffix of the first volume by typing a different number.
6. Click **Create** to effectively create and add the volumes to the storage pool (Figure 4-55).

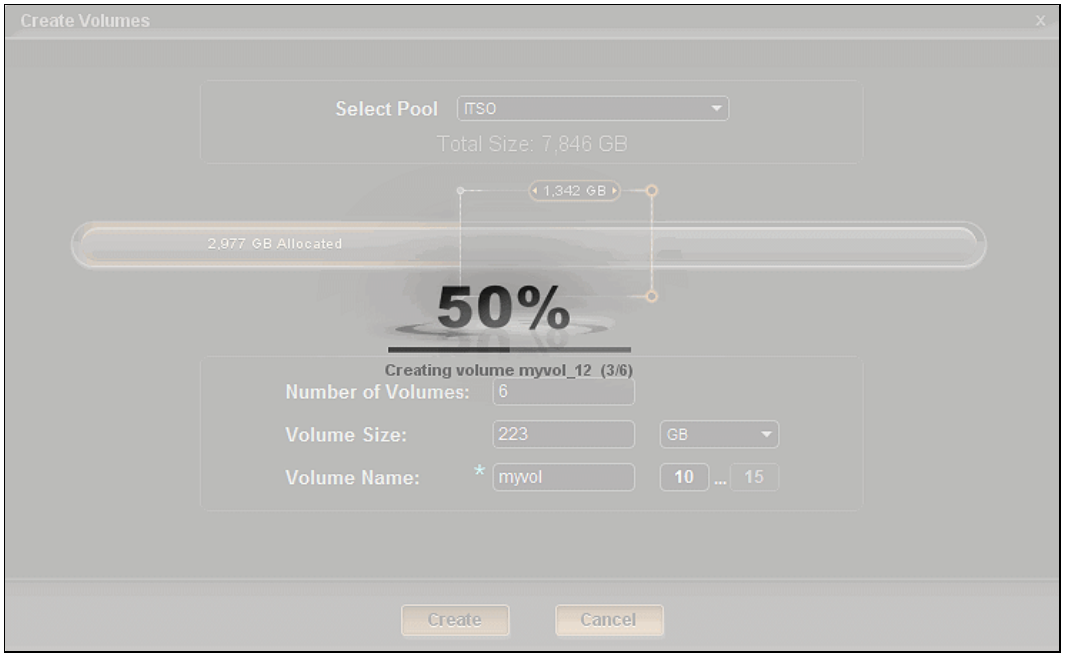


Figure 4-55 Volume creation progress indicator

After a volume is successfully added, its state is unlocked, meaning that write, format, and resize operations are permitted. The creation time of the volume is set to the current time and is never changed. Notice the volume name sequence in Figure 4-56.

Name	Size (GB)	Used (GB)	Consistency Group
ITSO		11% 7,846.0 GB Hard	
myvol_15	223 GB	0 GB	
myvol_14	223 GB	0 GB	
myvol_13	223 GB	0 GB	
myvol_12	223 GB	0 GB	
myvol_11	223 GB	0 GB	
myvol_10	223 GB	0 GB	

Figure 4-56 Volumes created

Resizing volumes

Resizing volumes is an operation that is similar to creating them. Only an unlocked volume can be resized. When you resize a volume, its size is specified as an integer multiple of 10^9 bytes. However, the actual new size of the volume is rounded up to the nearest valid size, which is an integer multiple of 17 GB.

Attention: The size of the volume can be decreased with the XIV Storage Management GUI. However, to avoid possible data loss, contact your IBM XIV Storage System support personnel if you need to decrease a volume size. (Mapped volume size cannot be decreased.)

Not all host operating systems support dynamic volume resizing. Consult with the host vendor documentation about whether your host OS supports resizing volumes dynamically and what restrictions and warnings there are.

The volume address space is extended (at the end of the existing volume) to reflect the increased size, and the additional capacity is logically formatted (that is, zeros are returned for all read commands).

When resizing a regular volume (not a writable snapshot), all storage space that is required to support the additional volume capacity is reserved (static allocation). This configuration guarantees the functions and integrity of the volume, regardless of the resource levels of the storage pool containing that volume.

Resizing a master volume does not change the size of its associated snapshots. These snapshots can still be used to restore their individual master volumes at their initial sizes.

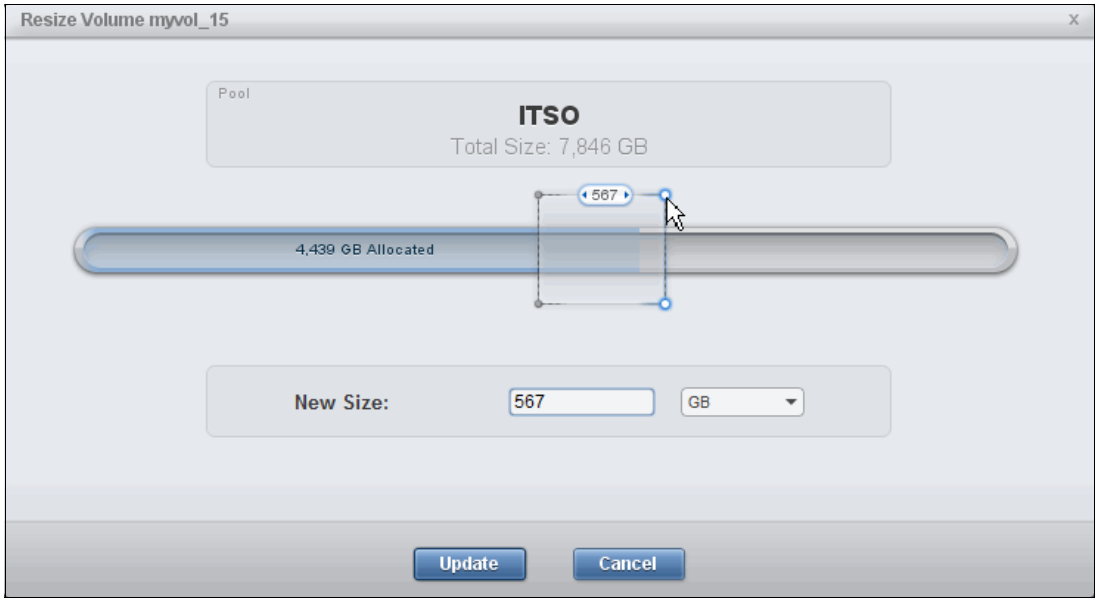


Figure 4-57 Resize an existing volume

To resize volumes using XIV Storage Management GUI, complete the following steps:

1. Right-click the row of the volume to be resized and select **Resize**.

   The total amount of storage is presented both textually and graphically. The amount that is already allocated by the other existing volumes is shown in blue. The amount that is free is shown in gray. The current size of the volume is displayed as a dotted outline box, around the storage pool gauge.

2. In the New Size field, type the new value, or drag the edge of the outline box.
3. Click **Update** to resize the volume.

Deleting volumes

Using the XIV Storage Management GUI, deleting a volume is as easy as creating one.

Important: After you delete a volume or a snapshot, all data stored on the volume is lost and cannot be restored.

All the storage space that was allocated (or reserved) for the volume or snapshot is freed and returned to its storage pool. The volume or snapshot is then removed from all the logical unit number (LUN) maps that contain mapping of this volume.

Deleting a volume deletes all the snapshots associated with this volume, even snapshots that are part of snapshot groups. A volume can be deleted even if the volume is in the lock state, but a volume cannot be deleted if the volume is mapped to a host or part of a consistency group.

To delete a volume or snapshot, complete the following steps:

1. Right-click the row of the volume to be deleted and select **Delete**.
2. Click to delete the volume.

Maintaining volumes

There are various other operations that can be issued on a volume. See “Menu option actions” on page 155 for more information.

The use of these operations is obvious, and you can initiate an operation by right-clicking one. These following operations are available:

- Format a volume:

  A formatted volume returns zeros as a response to any read command. The formatting of the volume is done logically, and no data is written to the physical storage space allocated for the volume. Consequently, the formatting action is performed instantly.

- Rename a volume:

  A volume can be renamed to a unique name in the system. A locked volume can also be renamed.

- Lock and unlock a volume:

  You can lock a volume so that hosts cannot write to it. A volume that is locked is write-protected, so that hosts can read the data stored on it, but they cannot change it. The volume displays then as a lock icon. In addition, a locked volume cannot be formatted or resized. In general, locking a volume prevents any operation (other than deletion) that changes the volume’s image.

  Lock and unlock: Master volumes are set to *unlocked* when they are created. Snapshots are set to *locked* when they are created.

- Create consistency groups:

  XIV Storage System enables a higher level of volume management provided by grouping volumes and snapshots into sets called *consistency groups*. This grouping is especially useful for cluster-specific volumes.

- Copy a volume:

  You can copy a source volume onto a target volume. All the data that was previously stored on the target volume is lost and cannot be restored.

- Use the snapshot functions:

  The XIV Storage System advanced snapshot feature has unique capabilities that enable the creation of a virtually unlimited number of copies of any volume, with no performance penalties.

- Map a volume:

  Although the storage system sees volumes and snapshots at the time of their creation, the volumes and snapshots are visible to the hosts only after the mapping procedure.

  For more information about mapping, see 4.6, “Host definition and mapping” on page 164.

- Change the SSD caching state:

  You can change the SSD caching state for selected volumes. By default, volumes follow the SSD caching system setting. With this option, you can manually override the default system setting for specific volumes. For more information, see 4.9, “Flash cache” on page 199.

4.5.2 Managing volumes with XIV Command-Line Interface

All of the operations explained in 4.5.1, “Managing volumes with the XIV Storage Management GUI” on page 153 can also be performed through the command-line interface. To get a list of all the volume-related commands, enter the following command in an XCLI Session:

```
help category=volume
```

Important: The commands shown in this section assume that you started an XCLI Session on the selected system, as described in “XCLI session features” on page 138. Replace the example name values to perform the commands.

Example 4-9 shows the output of the command.

Example 4-9 All the volume-related commands

```
Category  Name                 Description
volume    reservation_clear    Clears reservations of a volume.
volume    reservation_key_list Lists reservation keys.
volume    reservation_list     Lists volume reservations.
volume    vol_by_id            Prints the volume name according to its specified SCSI serial number.
volume    vol_copy             Copies a source volume onto a target volume.
volume    vol_create           Creates a new volume.
volume    vol_delete           Deletes a volume.
volume    vol_disable_vaai     Disables VAAI support for a specified volume
volume    vol_enable_vaai      Enables VAAI support for a specified volume
volume    vol_format           Formats a volume.
volume    vol_list             Lists all volumes or a specific one.
volume    vol_lock             Locks a volume so that it is read-only.
volume    vol_rename           Renames a volume.
volume    vol_resize           Resizes a volume.
volume    vol_unlock           Unlocks a volume, so that it is no longer read-only and can be written to.
```

To list the existing volumes in a system, run the following command:

```
vol_list pool="ITSO Pool"
```

For a list of existing volumes in a storage pool, run the following command:

```
vol_list pool=”ITSO Pool”
```

The result of this command is similar to the output shown in Figure 4-58.

```
Name      Size(GB)  Master Name  Consistency Group  Pool       Creator  Used Capacity
(GB)
myvol_10  223                                       ITSO Pool  itso     0
myvol_11  223                                       ITSO Pool  itso     0
myvol_12  223                                       ITSO Pool  itso     0
myvol_13  223                                       ITSO Pool  itso     0
myvol_14  223                                       ITSO Pool  itso     0
myvol_15  344                                       ITSO Pool  itso     0
```

Figure 4-58 vol_list command output

To find and list a specific volume by its SCSI ID (serial value), run the following command:

```
vol_by_id id=23
```

To create a volume, run the following command:

```
vol_create size=51 pool="ITSO Pool" vol="myvol_16"
```

The size can be specified either in gigabytes or in blocks (where each block is 512 bytes). If the size is specified in blocks, volumes are created in the exact size specified. If the size is specified in gigabytes, the actual volume size is rounded up to the nearest 17 GB multiple (making the actual size identical to the free space consumed by the volume). This rounding up prevents a situation where storage space is not fully used because of a gap between the free space used and the space available to the application.

The volume is logically formatted at creation time, which means that any read operation results in returning all zeros as a response. To format a volume, run the following command:

```
vol_format vol=”myvol_16”
```

All data stored on the volume is lost and unrecoverable. If you want to bypass the warning message, put **-y** directly after the XCLI command.

The following example shows how to resize one of the existing volumes:

```
vol_resize vol="myvol_16" size=103
```

Cannot shrink: If you attempt to decrease a volume’s size by using the XCLI, you receive a `CAN_NOT_SHRINK_VOLUME` message. Create a volume and migrate the data to the new volume.

To rename an existing volume, run the following command:

```
vol_rename new_name="myvol_17" vol="myvol_16"
```

To delete an existing created volume, run the following command:

```
vol_delete vol="myvol_17"
```

For details on how to manage volume SSD Caching through the XCLI, see 4.9.2, “Managing flash cache with XIV Command Line Interface” on page 204.

4.6 Host definition and mapping

Because the XIV Storage System can be attached to multiple, heterogeneous hosts, it is necessary to specify which particular host can access which specific logical drives in the XIV Storage System. Mappings must be defined between hosts and volumes in the XIV Storage System.

The XIV Storage System is able to manage single hosts or hosts grouped together in clusters.

4.6.1 Assigning LUNs to a host using the XIV Storage Management GUI

There are several steps that are required to define a new host and assign LUNs to it. One of the prerequisites is that volumes must be created in a storage pool.

Defining a host

To define a host, complete the following steps:

1. In the XIV Storage System main GUI window, hover the cursor over the Hosts and Clusters icon and select **Hosts and Clusters** (see Figure 4-59).

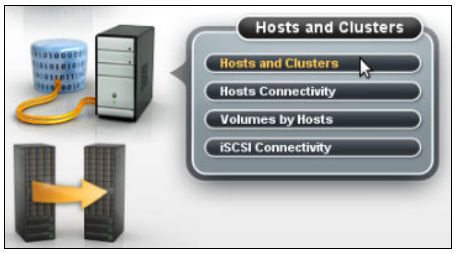


Figure 4-59 Hosts and Clusters menu

2. The Hosts window opens and shows a list of hosts (if any) that are already defined. To add a host or cluster, click either **Add Host** or **Add Cluster** in the menu bar (see Figure 4-60). In our example, we select **Add Host**. Add Host is used for a single host that is assigned a LUN or multiple LUNs. Add Cluster is used for a group of hosts that share a LUN or multiple LUNs.

Figure 4-60 Add new host

3. The Add Host window opens, as shown in Figure 4-61 on page 165. From the System / Cluster drop-down menu, choose a cluster for the host (choose **Standalone Hosts** for no cluster). Enter a Name for the host. Select a Type for the host (choose the default unless the host type is `hpux` or `zvm`).

Important: Do *not* change the host type of a host that has mapped volumes, because the volumes will be inaccessible to the host.

If you need to change the host type, remove all mappings first, change the host type, and remap the volumes afterward.

4. Optionally, if this host will use Challenge Handshake Authentication Protocol (CHAP) authentication for IP network Small Computer System Interface (iSCSI), enter values for CHAP Name and CHAP Secret.

Figure 4-61 Add host details

5. Repeat steps 2 and 3 to create additional hosts. In our scenario, we add another host that is called `ITSO_Win2008_iscsi`.
6. Host access to LUNs is granted depending on the host adapter ID. For a Fibre Channel (FC) connection, the host adapter ID is the FC host bus adapter (HBA) worldwide port name (WWPN). For an iSCSI connection, the host adapter ID is the host iSCSI qualified name (IQN). To add a WWPN or IQN to a host definition, right-click the host and select **Add Port** from the menu, as shown in Figure 4-62.

Figure 4-62 Add port to host definition

7. When the Add Port window shown in Figure 4-63 opens, select port type of FC or **iSCSI** from the drop-down menu.

   In this example, an FC host is defined. Add the WWPN for HBA1. If the host is correctly connected and has completed a port login to the storage area network (SAN) switch at least once, the WWPN is shown in the drop-down list box. Otherwise, you can manually enter the WWPN into the Port Name field. Adding ports from the drop-down list is less prone to error and is the best method. However, if hosts have not yet been connected to the SAN or zoned, then manually adding the WWPNs is the only option.

Figure 4-63 XIV Storage Management GUI example: Add FC port WWPN

   Repeat steps 5 and 6 to add the second HBA WWPN. Ports can be added in any order.

8. To add an iSCSI host, in the Add Port window, specify the port type as iSCSI and enter the IQN of the HBA as the iSCSI Name. See Figure 4-64.

Figure 4-64 XIV Storage Management GUI example: Add iSCSI port

9. The host is displayed with its ports in the Hosts window, as shown in Figure 4-65 on page 166.

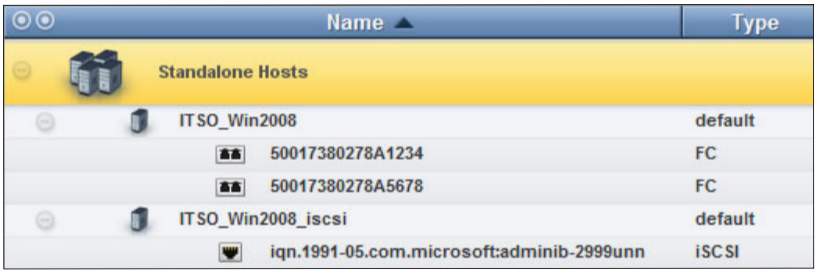


Figure 4-65 List of hosts and ports

In this example, the hosts `itso_Win2008` and `itso_Win2008_iscsi` are the same physical host, but they have been entered as separate entities so that when mapping LUNs, the FC and iSCSI protocols do not access the same LUNs.

Mapping logical unit numbers to a host

The final configuration step is to map logical unit numbers (LUNs) to the host. To do this task, complete the following steps:

1. While still in the Hosts and Clusters configuration pane, right-click the host to which the volume will be mapped and select **Modify LUN Mapping** from the menu (see Figure 4-66).

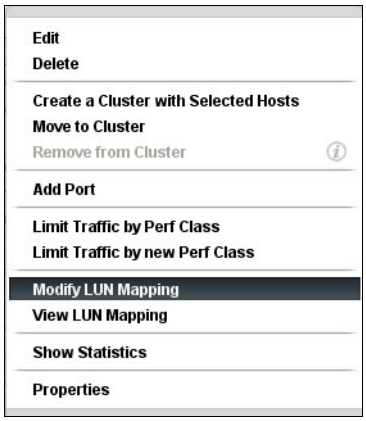


Figure 4-66 Map LUN to host

2. When the LUN Mapping for Host window opens (Figure 4-67), select an available volume from the left pane.

   The XIV Storage Management GUI suggests a LUN ID to which to map the volume, which can be changed to meet your requirements.

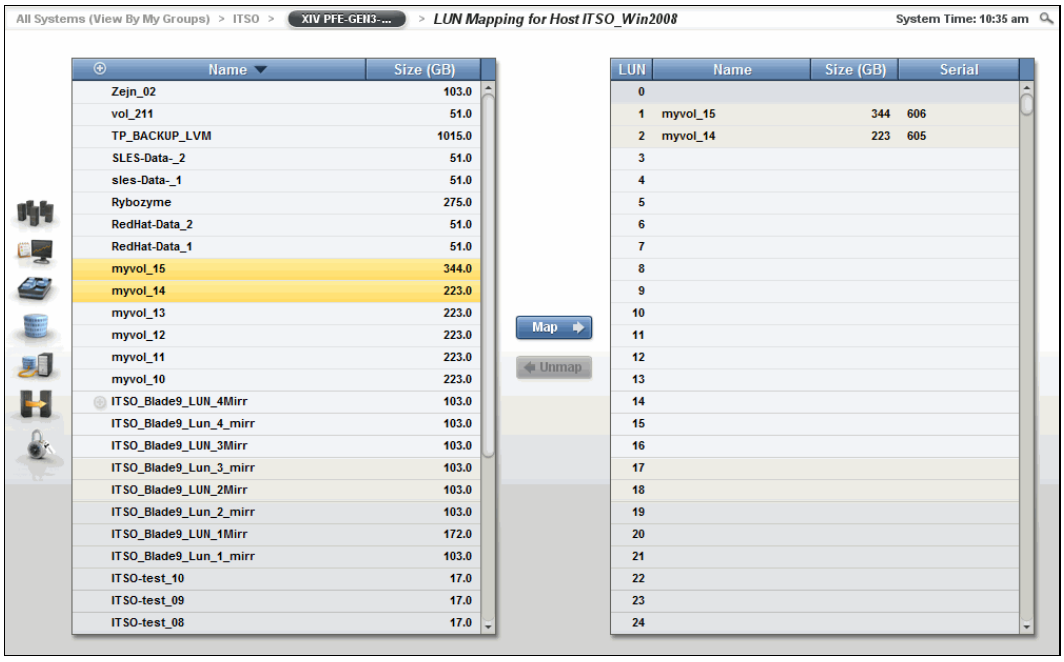


Figure 4-67 Map FC volume to FC host

3. Click **Map** and the volume is assigned immediately.

   There is no difference in mapping a volume to an FC or iSCSI host in the XIV Storage Management GUI LUN Mapping for Host view.

4. To complete this example, start the host server and check connectivity. The XIV Storage System has a real-time connectivity status overview.

5. Click **Hosts and Clusters** → **Hosts Connectivity** to access the connectivity status. See Figure 4-68.

Figure 4-68 Hosts Connectivity

The Host Connectivity window opens. In the example, the `ITSO_Win2008` and `ITSO_Win2008_iscsi` host definitions, which are shown in Figure 4-69, are shown with their corresponding FC and iSCSI connections to the XIV Storage System.

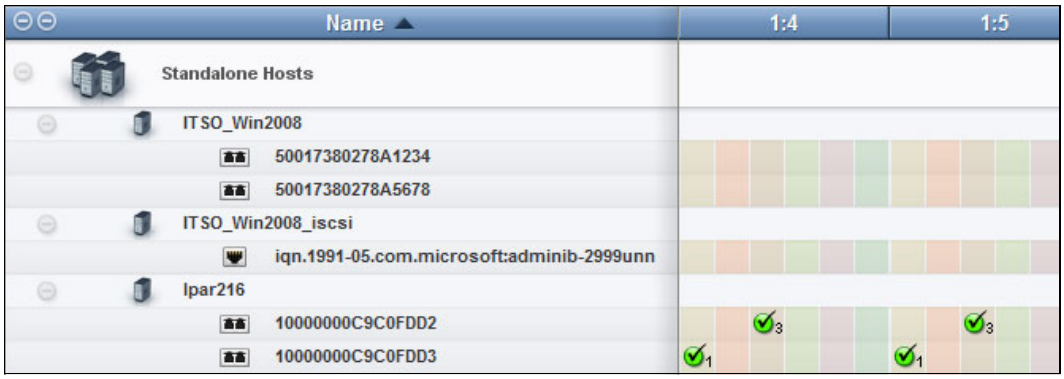


Figure 4-69 XIV Storage Management GUI example: Host connectivity matrix

Tip: The Hosts Connectivity view contains icons that indicate the port type (iSCSI or FC) for each port.

4.6.2 Assigning logical unit numbers to a host by using the XCLI

There are several required steps to define a new host and assign LUNs to it. One of the prerequisites is that volumes are created in a storage pool.

Defining a new host

To use the XCLI to define a new host, complete the following steps:

1. Create a host definition for your FC and iSCSI hosts by running **`host_define`**, as shown in Example 4-10.

Example 4-10 XCLI Session example: Create host definition

```
>>host_define host=ITSO_Win2008
```

```
Command executed successfully.

>>host_define host=ITSO_Win2008_iscsi
Command executed successfully.
```

Host access to LUNs is granted depending on the host adapter ID. For an FC connection, the host adapter ID is the FC HBA WWPN. For an iSCSI connection, the host adapter ID is the IQN of the host.

2. Add the WWPN of the FC host for HBA1 and HBA2 by running **host_add_port** with the **fcaddress** parameter, as shown in Example 4-11.

Example 4-11 Create FC port and add to host definition

```
>> host_add_port host=ITSO_Win2008 fcaddress=10000000C9831FA2
Command executed successfully.
>> host_add_port host=ITSO_Win2008 fcaddress=10000000C9831FA3
Command executed successfully.
```

3. Add the IQN of the iSCSI host, as shown in Example 4-12 on page 169. This command is the same **host_add_port** command, but with the **iscsi_name** parameter.

Example 4-12 Create iSCSI port and add to the host definition

```
>> host_add_port host=ITSO_Win2008_iscsi
iscsi_name=iqn.1991-05.com.microsoft:win-gj5e8kr49ee.itso.storage.ibm.com
Command executed successfully
```

Mapping logical unit numbers to a host

The final configuration step is to map LUNs to the host definition. Complete the following steps:

1. Map LUNs to the host definition by using the commands shown in Example 4-13. For a cluster, the volumes are mapped to the cluster host definition. There is no difference for FC or iSCSI mapping to a host.

Example 4-13 XCLI example: Map volumes to hosts

```
>> map_vol host=ITSO_Win2008 vol=myvol_14 lun=1
Command executed successfully.

>> map_vol host=ITSO_Win2008 vol=myvol_15 lun=2
Command executed successfully.

>> map_vol host=ITSO_Win2008_iscsi vol=myvol_10 lun=1
Command executed successfully.
```

2. To complete the example, start the server and check the host connectivity status from the XIV Storage System point of view. Example 4-14 shows the output for both hosts.

Example 4-14 XCLI example: Check host connectivity

```
>> host_connectivity_list host=ITSO_Win2008
Host          Host Port         Module      Local FC port  Local iSCSI port  Type
ITSO_Win2008  10000000C9831FA2  1:Module:9  1:FC_Port:9:2                    FC
ITSO_Win2008  10000000C9831FA3  1:Module:5  1:FC_Port:5:2                    FC
>> host_connectivity_list host=ITSO_Win2008_iscsi
Host               Host Port
Module     Local FC port  Type
ITSO_Win2008_iscsi  iqn.1991-05.com.microsoft:win-gj5e8kr49ee.itso.storage.ibm.com
1:Module:6               SCSI
```

```
ITSO_Win2008_iscsi  iqn.1991-05.com.microsoft:win-gj5e8kr49ee.itso.storage.ibm.com
1:Module:5                iSCSI
```

In Example 4-14, there are two paths per host FC HBA and two paths for the single Ethernet port that was configured.

See *XIV Storage System: Host Attachment and Interoperability*, SG24-7904 for details related to host definitions and volume mapping.

4.7 QoS feature

The Quality of Service (QoS) feature allows the XIV Storage System to deliver different service levels to hosts that are connected to the same XIV system.

The QoS feature favors performance of critical business applications that run concurrently with noncritical applications. Because the XIV disk and cache are shared among all applications and all hosts are attached to the same resources, division of these resources among both critical and noncritical applications might have an unintended adverse performance effect on critical applications. QoS can address this by limiting the rate, based on bandwidth and IOPS, for non-critical applications. Limiting performance resources for non-critical applications means that the remaining resources are available without limitation for the business-critical applications.

The QoS feature is managed through the definition of performance classes and then associating hosts with a performance class. The feature was extended in the XIV Storage Software Version 11.5 and can also be set by XIV domains and XIV storage pools. Each performance class is now implicitly one of two types: host type or pool/domain type.

The QoS feature possibilities and limitations can be summarized as follows:

- Up to 500 performance classes are configurable.
- QoS is applicable to host, domain, pool and restricted combinations of these entities. For instance, hosts cannot be specified for a performance class that already contains a domain or pool (see Figure 4-73 on page 172).
- Limits can be defined as *Total*, meaning for XIV system as a whole, or *Per Interface*.
- Limits are specified as IOPS or bandwidth.
- Limit calculation is based on preferred practices for setup and zoning.

  The limited I/O processes are expected to always come through all active XIV interface nodes (equal to active interface modules). For example, on a 9module partial rack XIV, where 4 interface modules are active, the total I/O or bandwidth rate would be divided by 4 (the number active interface modules). If a limit total of 3,000 I/Os is specified, it would result to a limitation of 750 I/Os per interface module.

  In addition, in the case of the 9-module XIV, if the limited I/Os are coming through only two of the four interface modules (as a result of the SAN zoning), the effective limitation will be 2 x 750 I/Os = 1,500 I/Os rather than the expected 3,000 I/O limitation.

  The calculation-related information is visible in the GUI, as shown in Figure 4-70, for a 9- module XIV Storage System (partial rack configuration).

Figure 4-70 QoS, calculation based on active XIV interface modules (9-module XIV Storage System)

A similar calculation takes place for the bandwidth limit.

Note: If more than one host, domain, or pool is added to a performance class, all hosts, domains, or pools in this performance class share the limitations defined on that performance class. For example, if two or more entities are added to a 10,000 IOPS performance class, the total number of all contained entities IOPS is limited to 10,000. Therefore, it is a good practice to create one performance class per domain and one performance class per pool.

4.7.1 Managing QoS

The management of the QoS feature is quite easy. To define a performance class, click the **QoS Performance Class** option from the Monitor menu (Figure 4-71 on page 171).

Figure 4-71 QoS: menu location

To configure a performance class, complete the following steps:

1. Select the QoS Performance Class view from the Monitor menu, and click **Add Performance Class** on the toolbar. The dialog window shown in Figure 4-72 opens.

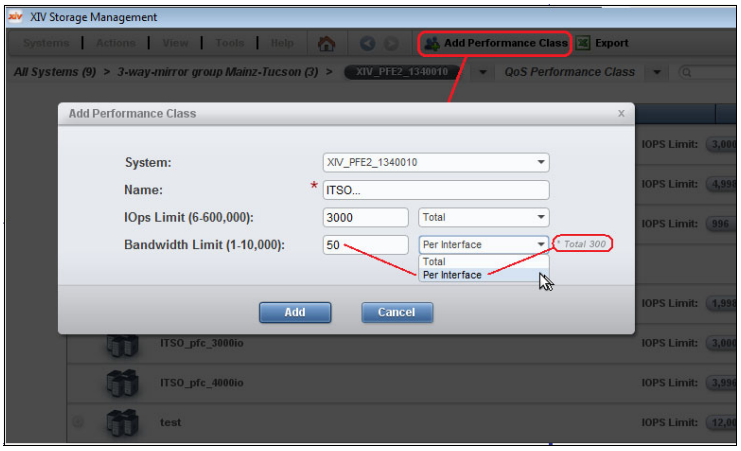


Figure 4-72 QoS: Add Performance Class window

2. Enter a suitable performance class name and an IOPS limit, a bandwidth limit, or a combination of both, based on business needs. There are two choices when entering the limitation settings: *Total* or *Per Interface*.

   The Total intended limitation settings depend on the number of interface modules being used, as previously described.

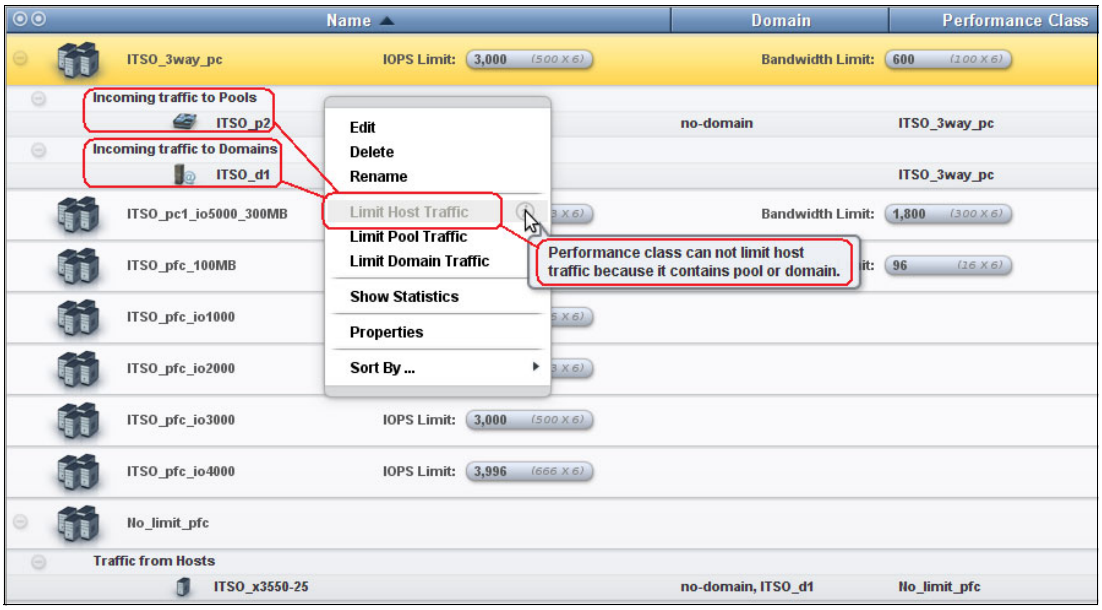


Figure 4-73 QoS limitation: Performance classes cannot contain hosts along with domains or pools

Conversely, QoS settings can also be specified on a per-interface module level. Taking into account the information presented earlier about a fully populated system, the maximum QoS IOPS value at the per interface level is 100,000 IOPS, with a minimum of 1 IOPS. The maximum total QoS bandwidth value at the per interface level is 10,000 MB/s, with a minimum of 1 MB/s. To adapt these to the XIV total limits for a performance class, multiply by the number of active interface modules. For example, for a

fully populated XIV system with 15 modules and 6 active interface modules, multiply by 6. For a partial system with 9 modules and 4 active interface modules, multiply by 4.

Performance class creation using the XCLI: Using the XCLI enables you to create an unlimited performance class (which is not possible from the GUI), as shown in Example 4-15. The unlimited performance class is useful when using the drag-and-drop function of the GUI for dynamic adjustments based on business needs.

Example 4-15 QoS: Creating a 'no limit' performance class with XCLI

```
XIV_PFE2_1340010>>perf_class_create perf_class=No_limit_pfc
Command executed successfully.
XIV_PFE2_1340010>>perf_class_list
Performance class       Max IO rate(IOPS)  Max BW rate(MB/s)
No_limit_pfc            0                  0
XIV_PFE2_1340010>>
```

XCLI commands related to performance class handling are self explanatory and listed below in Example 4-16 on page 173.

Example 4-16 QoS: XCLI performance class command set

```
XIV_PFE2_1340010>>help search=perf_class
Category    Name                     Description
system      perf_class_list          Lists Performance Classes
system      perf_class_create        Creates a Performance Class
system      perf_class_delete        Deletes a Performance Class
system      perf_class_rename        Renames a Performance Class
system      perf_class_set_rate      Sets the rate for  a Performance Class
system      perf_class_add_host      Adds a host to a Performance Class
system      perf_class_remove_host   Removes a host from its Performance Class
system      perf_class_add_pool      Adds a pool to a Performance Class
system      perf_class_remove_pool   Removes a pool from its Performance Class
system      perf_class_add_domain    Adds a Domain to a Performance Class.
system      perf_class_remove_domain Removes a Domain from its Performance Class.
XIV_PFE2_1340010>>
```

Figure 4-74 shows the GUI overview of performance classes, where hosts in the Unlimited Group (`No_limit_pfc`) can be dragged to the QoS class. In addition, both IOPS and bandwidth limits can be adjusted by right-clicking either setting and clicking **Edit**.

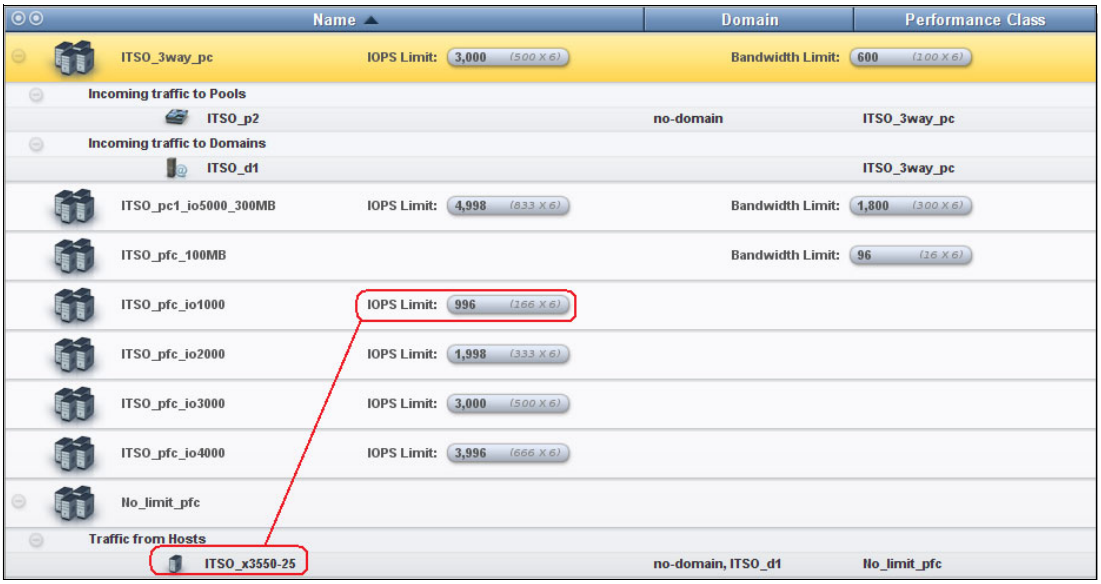


Figure 4-74 QoS: GUI, drag and drop of a host between performance classes

To export all the details shown in the QoS GUI overview, click the **Export** icon in the toolbar at the top of the GUI window. This function creates a `*.csv` file that contains the information about each defined performance classes.

As you can see, the XIV GUI offers an intuitive method for managing QoS delivery in IBM XIV systems.

The XCLI also allows you to schedule QoS features through scripts. The business requirements can be accommodated dynamically by modifying the QoS classes to ensure that the wanted level of performance is achieved on time.

Important: When using the QoS performance class feature, XIV consistently limits throughput to the specified QoS limits, regardless of the host, domain, or pool I/O demand.

4.8 Multitenancy

XIV multitenancy brings flexibility and simplicity to management of tenant data and storage resources across multiple XIV systems. Multitenency provides several advantages:

- Secure division and isolation of XIV Storage resources among numerous tenants
- Simple, quick delegation of administration tasks and role-based permissions
- Simple, rapid deployment without the need for extensive planning and tuning, as well as field-upgradability

In this section, we cover the details of domain administration using the XIV Storage System management software.

4.8.1 Domains

As discussed in 2.4.3, “Multitenancy” on page 30, a domain is a logical subset of the XIV system. Just as an XIV Storage System has hard and soft capacities, so does a domain. When you create a domain, you define the hard and soft capacities available to that domain, and those parameters define the aggregate hard and soft capacities that can be consumed and allocated by the pools within the domain. For more information about the hard and soft capacity concepts, see 2.7, “Capacity allocation and thin provisioning” on page 37.

Global administrators and domain administrators

To facilitate the creation and subsequent management of domains, the XIV Storage System 11.5 code release introduces the concepts of *global administrators* and *domain administrators*.

- A global administrator is a user who is not associated with any domain and that has access rights (as implied by the user’s role, such as, storageadmin) to all of the entities that are not uniquely associated with any domain.
- Conversely, a domain administrator is a user who is associated with one or more domains and has access rights (as implied by the user’s role, such as, storageadmin) to the entities uniquely associated with his or her associated domains.

4.8.2 Managing domains with the XIV Storage Management GUI

In this section, we describe how to manage domains using the XIV Storage Management GUI. The topic we cover include:

- Creating domains
- Viewing and editing domain properties
- Managing domain associations
- Viewing domain associations
- Limiting domains by performance class (QoS)
- Deleting domains

Managing domains with the XIV Storage Management GUI is fairly simple and intuitive. The domains page can be reached either through the **View** → **All Systems** → **Domains** option in the menu bar or from the **Domains** option in the **Systems** function icon on the left, as shown in Figure 4-75.

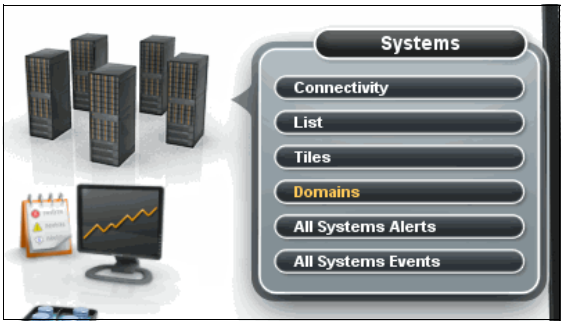


Figure 4-75 Opening the Domains page from the left menu

The domains page displays some or all of the domains currently defined on the system you are managing, as shown in Figure 4-76.

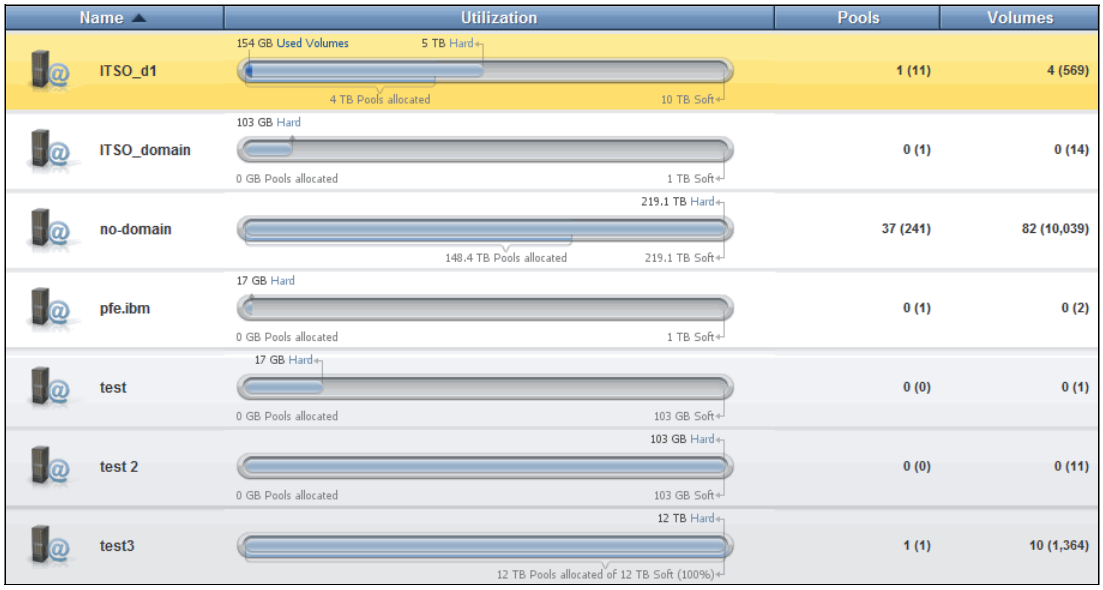


Figure 4-76 The Domains view

Note: The specific domains displayed in the Domains view depend upon the domain associations of the logged in user. The global storage administrator sees all domains defined on the XIV system. A user who is associated with one or more domains sees only those domains. For details on the distinction between global administrators and domain administrators, see “Global administrator and domain administrator” on page 223.

Tip: To filter the contents of the Domains view or any other view to show only items related to a specific domain, use the filter field in the menu bar to specify the domain, using the format domain:*domainName*, as shown in Figure 4-77

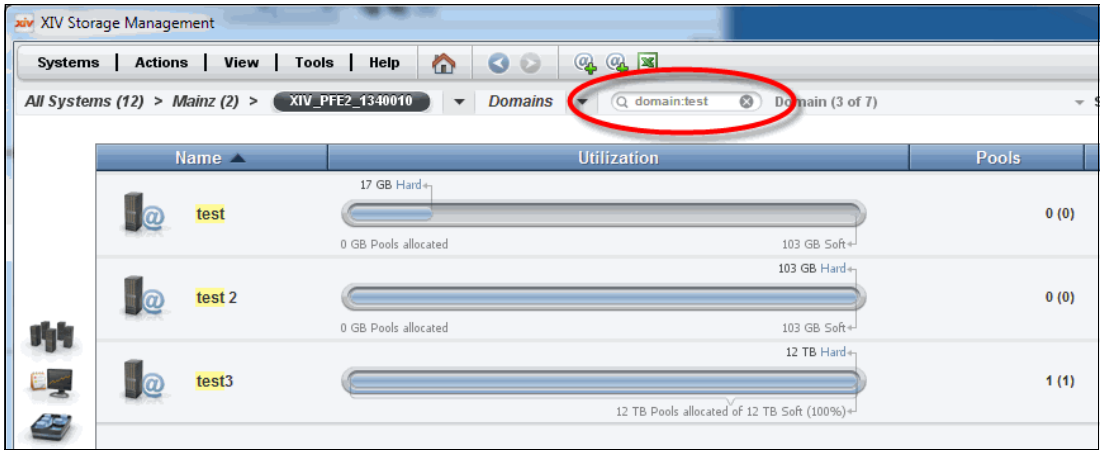


Figure 4-77 Filtering a GUI view by domain

The Domains view displays a table of domains in the system. The default view displays the domain name, utilization information, number of pools and number of volumes for each pool. This view gives the administrator a quick grasp and general overview of essential information about the system pools. The Domains view can be customized to show or hide specific columns. The Name and Utilization columns are required. The following optional column values are available:

- System
- Hard Capacity
- Free Hard
- Soft Capacity
- Free Soft
- Used Capacity
- Pools
- Free Pools
- Volumes
- Free Volumes
- Consistency Groups
- Free Consistency Groups
- Mirrors
- Free Mirrors
- Data Migrations
- Free Data Migrations
- LDAP Id
- Performance Class
- Allow SSD Caching

The capacity consumption by volumes and snapshots within the pools assigned to a particular domain is indicated by various colors. These are the default threshold values:

Blue	Capacity consumption below 80%
Yellow	Capacity consumption above 80%
Orange	Capacity consumption of over 90%
Red	Storage pool has depleted hard capacity

The name, the size, and separated segments are labeled appropriately.

Note: As mentioned in “Pool alert thresholds” on page 149, Version 11.5 introduces the ability to set pool alert thresholds on a per-pool basis, and consequently thresholds can be controlled on a per domain basis.

Creating domains

The creation of a domain with the XIV GUI is a simple process. To create a domain by using the GUI, follow these steps:

1. Click **Actions** → **Create Domain** from the menu bar, as shown in Figure 4-78.

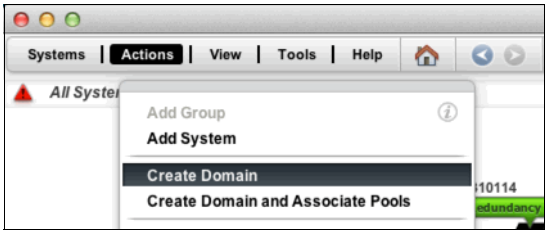


Figure 4-78 Creating a domain from the Actions menu in the menu bar

Alternatively, you can click **Create Domain** in the menu bar on the Domains view page, as shown in Figure 4-79.

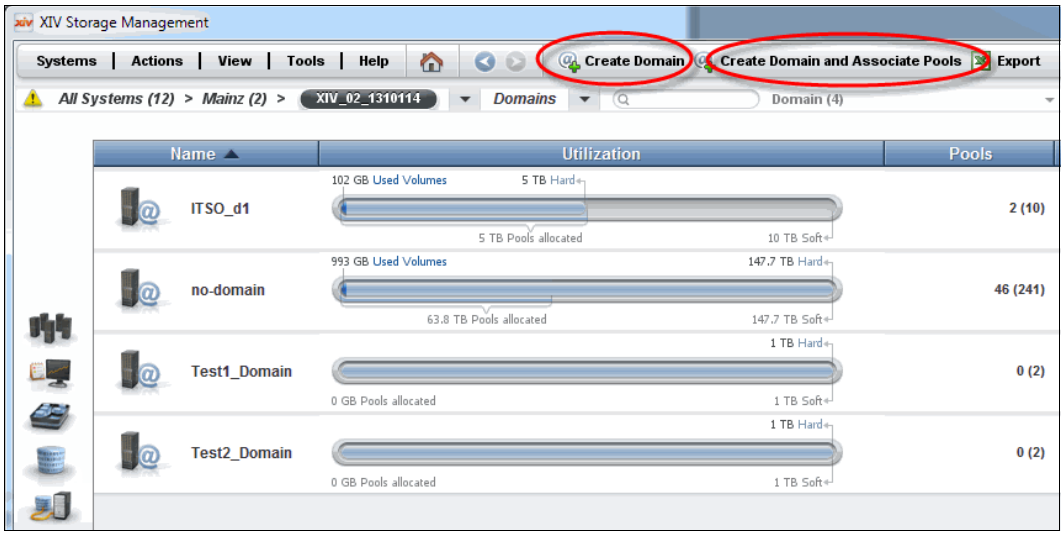


Figure 4-79 Creating a domain using the Create Domain label in the menu bar

Note: As shown in Figure 4-78 on page 177 and Figure 4-79, you have the option to associate existing pools to a domain when creating it. If you choose not to associate any existing pools to the domain upon creation, you can still associate existing pools to that domain at a later time.

Important: Only a global storage administrator, as defined in “Domain creation” on page 31, can create a domain and have the option to associate existing storage pools with a domain.

2. If you chose **Create Domain** in step 1 on page 177, you will see the Create Domain screen as shown in Figure 3Figure 4-80. The screen loads on the Capacity tab by dedault.

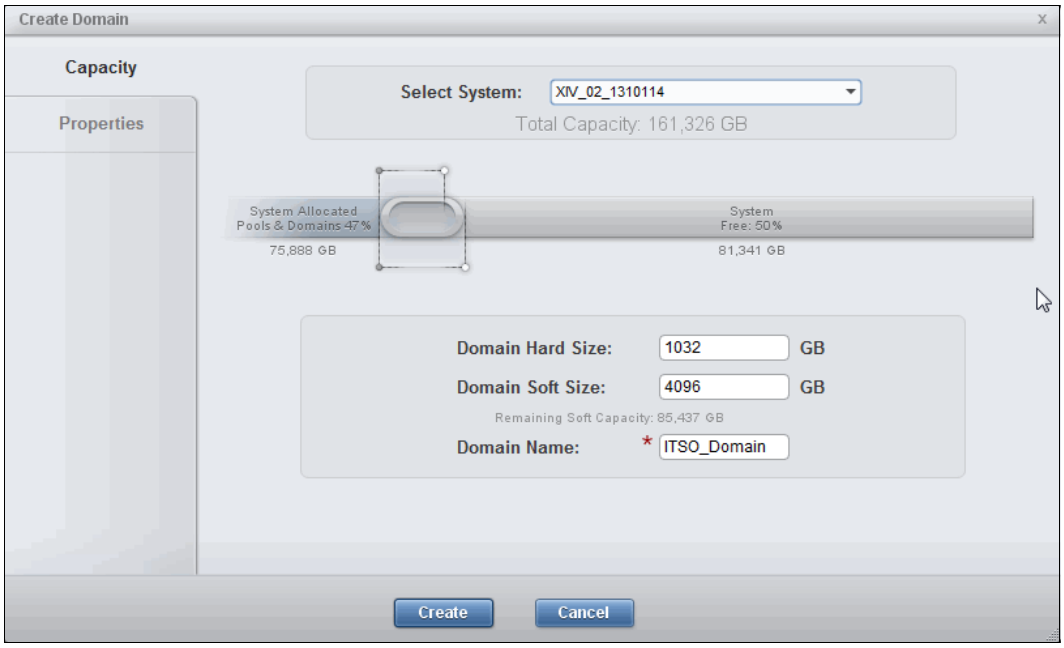


Figure 4-80 The Capacity tab of the Create Domain window in the XIV GUI

3. On the Capacity tab, select the XIV system on which you want to create the domain, specify the Domain Hard and Soft Sizes, and enter a unique Domain Name.
4. Then, click the Properties tab to specify the additional properties of the domain, as shown in Figure 4-81. Note, the GUI automatically displays the available and suggested numbers for the numerical properties.

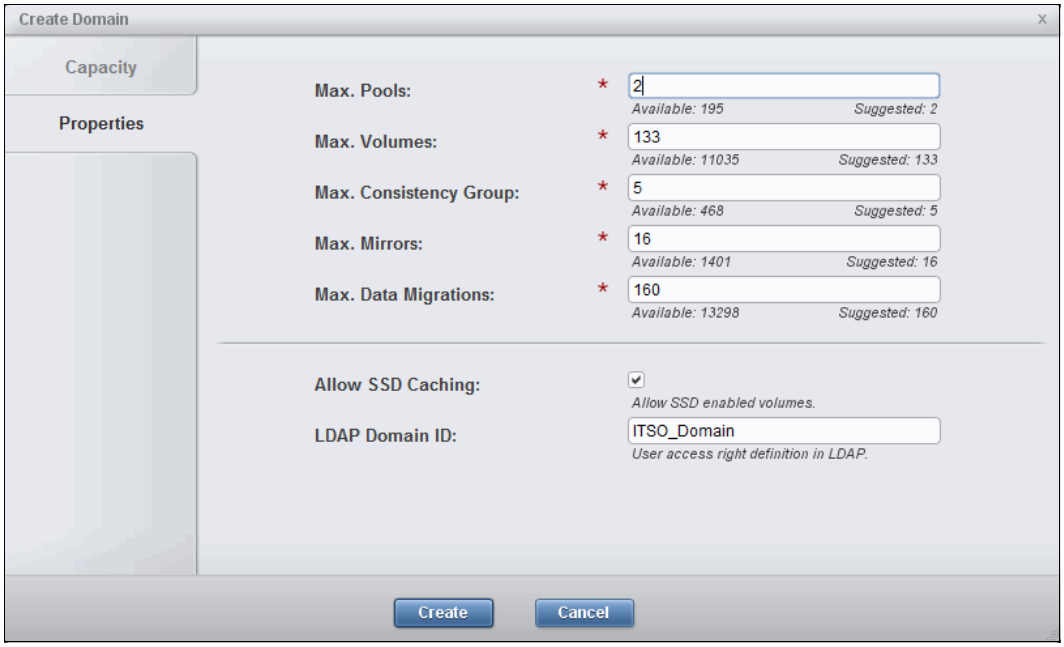


Figure 4-81 The Properties tab of the Create Domain screen in the XIV GUI

These properties are required:

Max. Pools	The maximum number of pools that can be associated with this domain
Max. Volumes	The maximum number of volumes that can be contained within this domain
Max. Consistency Groups	The maximum number of consistency groups that can be defined within this domain
Max. Mirrors	The maximum number of remote mirrors that can be defined within this domain
Max. Data Migrations	The maximum number of data migrations that can be defined within this domain
Allow SSD Caching	True or False, whether to allow SSD caching to be enabled on the volumes contained within this domain
LDAP Domain ID	The ID attribute of the domain to use when configuring multitenancy with LDAP authentication

Note: By default, the LDAP Domain ID field is populated with the value of the Domain Name when Properties tab loads. Optionally, you can modify this value. If you clear this field, the LDAP Domain ID value of the domain will automatically be set to the value of the Domain Name. For more information about configuring multitenancy with LDAP authentication, see 4.8.4, “Domains and LDAP authentication” on page 195.

5. After completing the fields in the Properties tab, click **Create** to create the domain.

   If you chose **Create Domain and Associate Pools** in step 1 on page 177, you will see the Create Domain and Associate Pools screen shown in Figure 4-82.

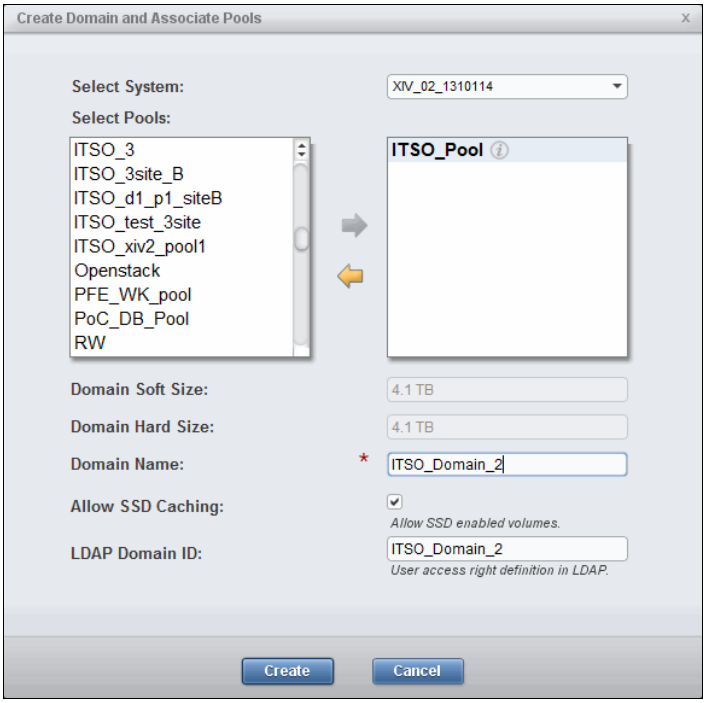


Figure 4-82 The Create Domain and Associate Pools screen in the XIV GUI

On this screen, select the XIV system on which you want to create the domain, specify the pools you want to associate with the domain, and enter a unique Domain Name. Also indicate whether to allow SSD caching on the volumes within this domain, and, optionally, modify the LDAP Domain ID to use when configuring multitenancy with LDAP authentication.

6. After completing the fields in the Create Domain and Associate Pools tabs, click **Create** to create the domain.

   If you choose to create a domain and associate it with existing pools, the domain will automatically be created with the minimum resources (those shown in the Properties tab in Figure 4-81 on page 179) to support the selected pools, as indicated in Figure 4-83.

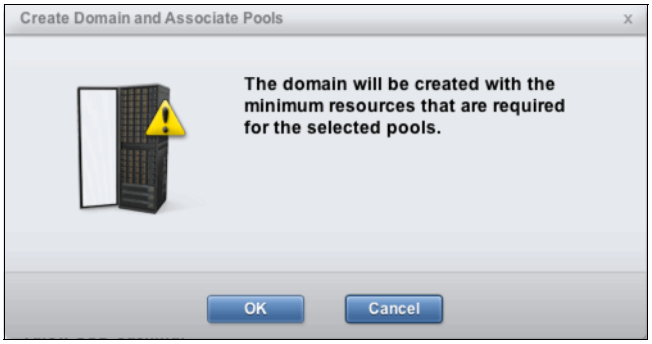

Create Domain and Associate Pools
The domain will be created with the minimum resources that are required for the selected pools.
OK
Cancel

Figure 4-83 Alert when creating a domain and associating pools

7. You can now verify that the domain was created from the Domains view, as show in Figure 4-84.

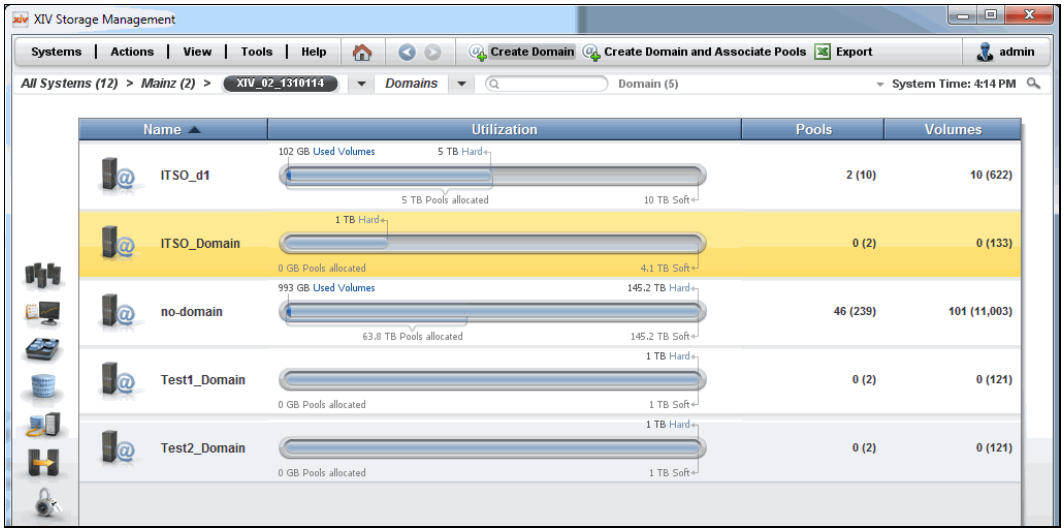


Figure 4-84 A successfully created domain shown in the Domains view in the XIV GUI

Viewing and editing domain properties

After you have created a domain in the XIV Storage System, it is easy to view or edit the properties of that domain using the XIV GUI.

To view the properties of a domain using the GUI, right-click the domain in the Domains view, and click Properties, as shown in Figure 4-85 on page 182.

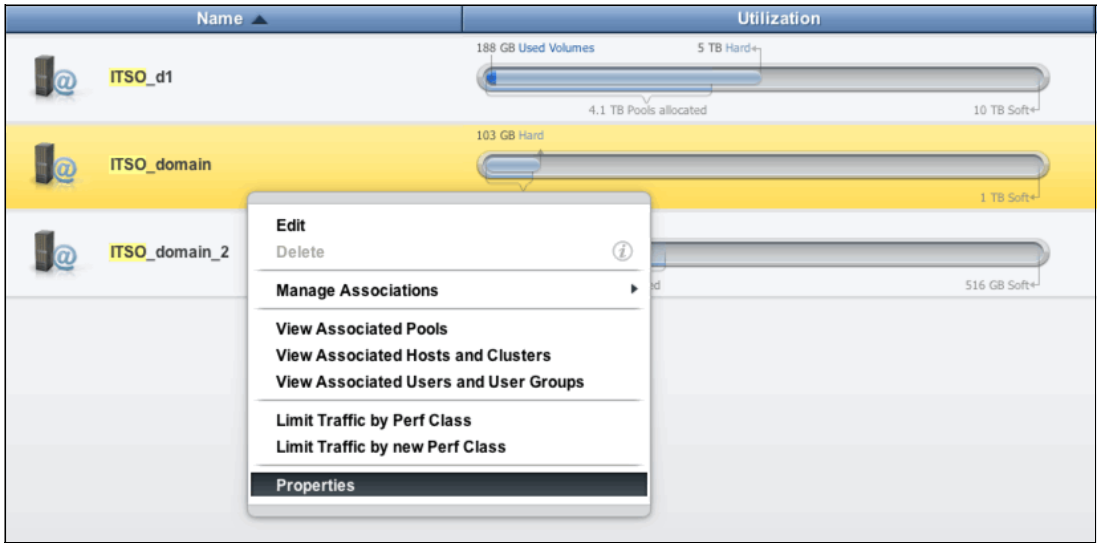


Figure 4-85 Accessing the Properties pane of a domain in the XIV GUI

This opens the Domain Properties pane for the selected domain, which displays the properties of the domain, as shown in Figure 4-86.

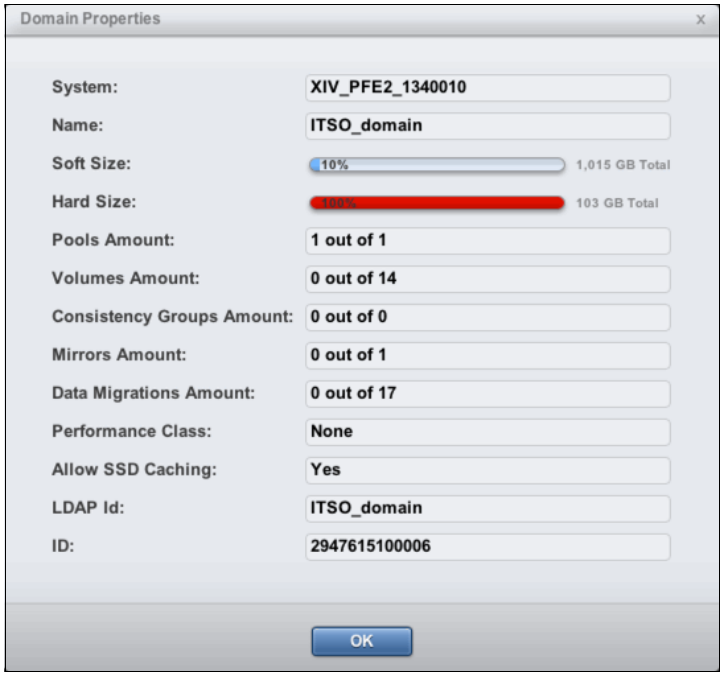

Domain Properties
System: XIV_PFE2_1340010
Name: ITSO_domain
Soft Size: 10% 1,015 GB Total
Hard Size: 100% 103 GB Total
Pools Amount: 1 out of 1
Volumes Amount: 0 out of 14
Consistency Groups Amount: 0 out of 0
Mirrors Amount: 0 out of 1
Data Migrations Amount: 0 out of 17
Performance Class: None
Allow SSD Caching: Yes
LDAP Id: ITSO_domain
ID: 2947615100006
OK

Figure 4-86 The Domain Properties pane in the XIV GUI

To edit the properties of a domain using the GUI, right-click the domain in the Domains view, and click **Edit**, as shown in Figure 4-87 on page 183.

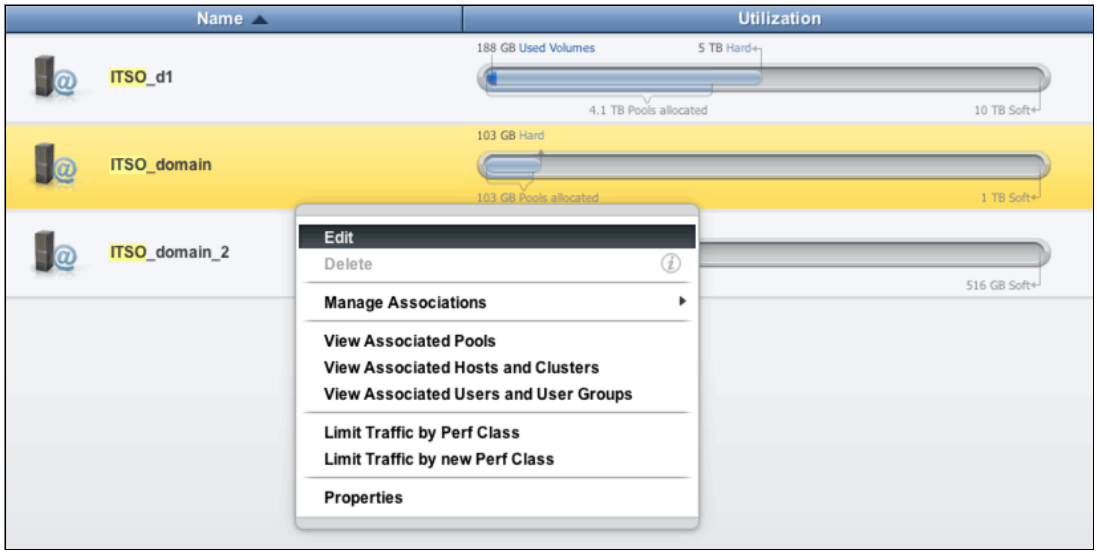


Figure 4-87 Accessing the Edit pane of a domain in the XIV GUI

This opens the Edit Domain window for the selected domain, which displays and allows you to edit the properties of the domain, as shown in Figure 4-88.

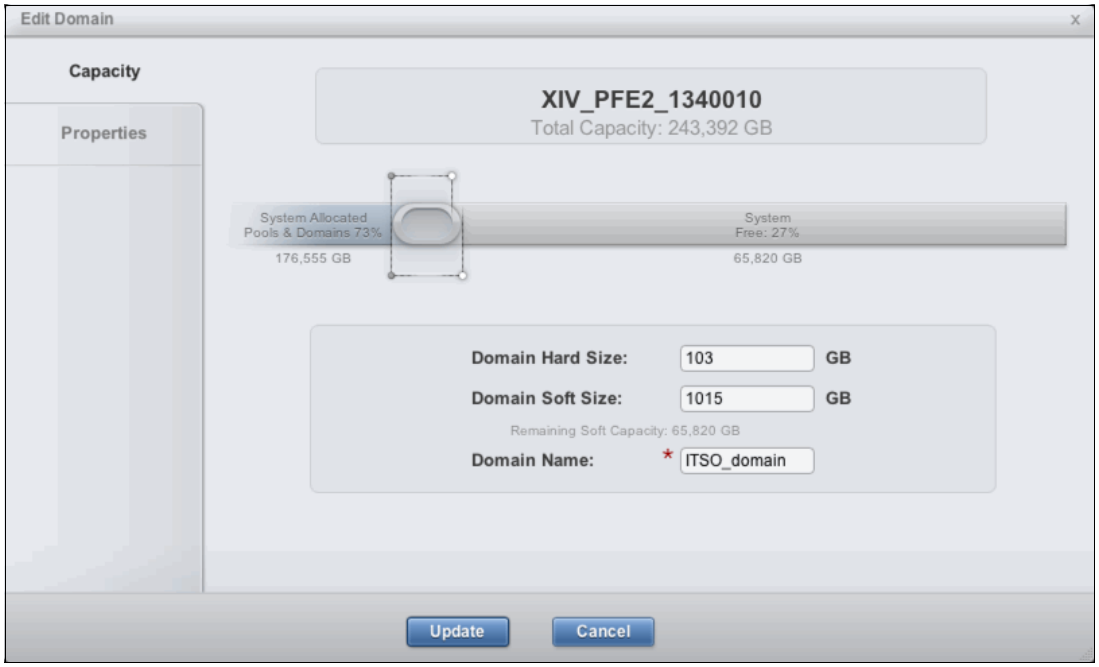


Figure 4-88 The Edit Domain window in the XIV GUI

In the Edit Domain window, you can modify the properties of the domain as required. For more information about the properties of a domain, see “The name, the size, and separated segments are labeled appropriately.” on page 177.

Managing domain associations

After a domain has been created, a global administrator with a Storage Administrator role can manage the system resources associated with that domain. The resource associations that such a user can manage include:

- Pools
- Hosts
- Clusters
- Targets
- Users*
- User Groups

***Note:** Domain user associations can only be managed by the Global Storage Administrator through the XIV GUI when local authentication is in use. When LDAP authentication is in use (enabled), domain user associations must be managed directly within the chosen LDAP solution.

Also, with LDAP enabled, a global administrator can associate user groups to a domain, but cannot disassociate user groups from a domain with LDAP active.

Further, a domain administrator with a Storage Administrator role has the ability to manage domain user and user group associations. However, such a user has no ability to manage other domain resource associations.

Note: When local authentication is active, only local users that are *not* associated with a local group can be associated with a domain. To associate a local user that is a member of a local group with a domain, you must associate the user group with the domain.

For more information about domain administrators in XIV, see “Global administrator and domain administrator” on page 223.

Managing domain pool associations

To manage pool associations with a domain, right-click the domain in the Domains view, and select **Manage Associations** → **Pools**, as shown in Figure 4-89.

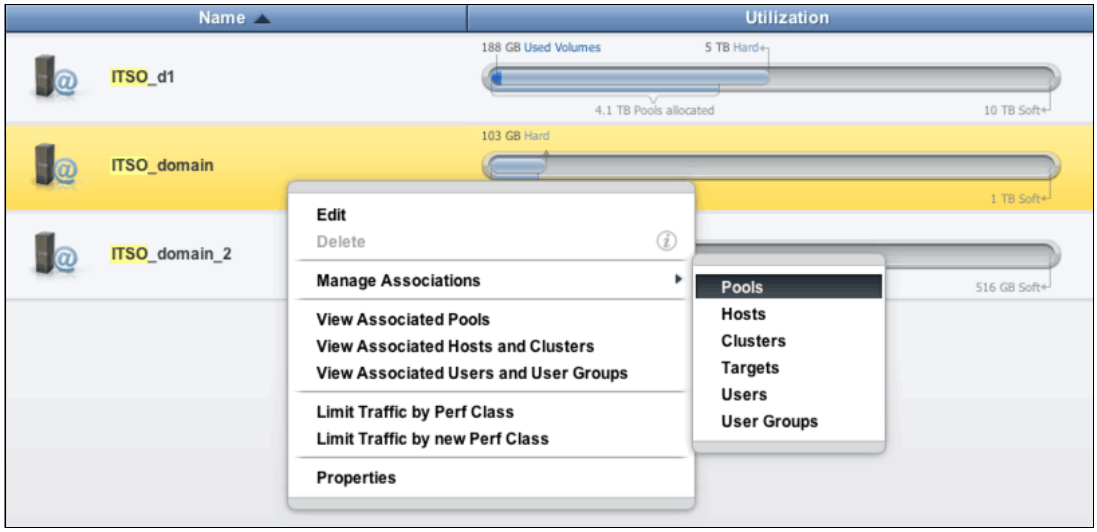


Figure 4-89 Accessing the Manage Domain’s Associations screen Pools tab in the XIV GUI

In the Manage Domain's Associations screen that appears (Figure 4-90), select the pools and use the arrows to associate or remove the pools from the domain. Click **Update** to apply the changes.

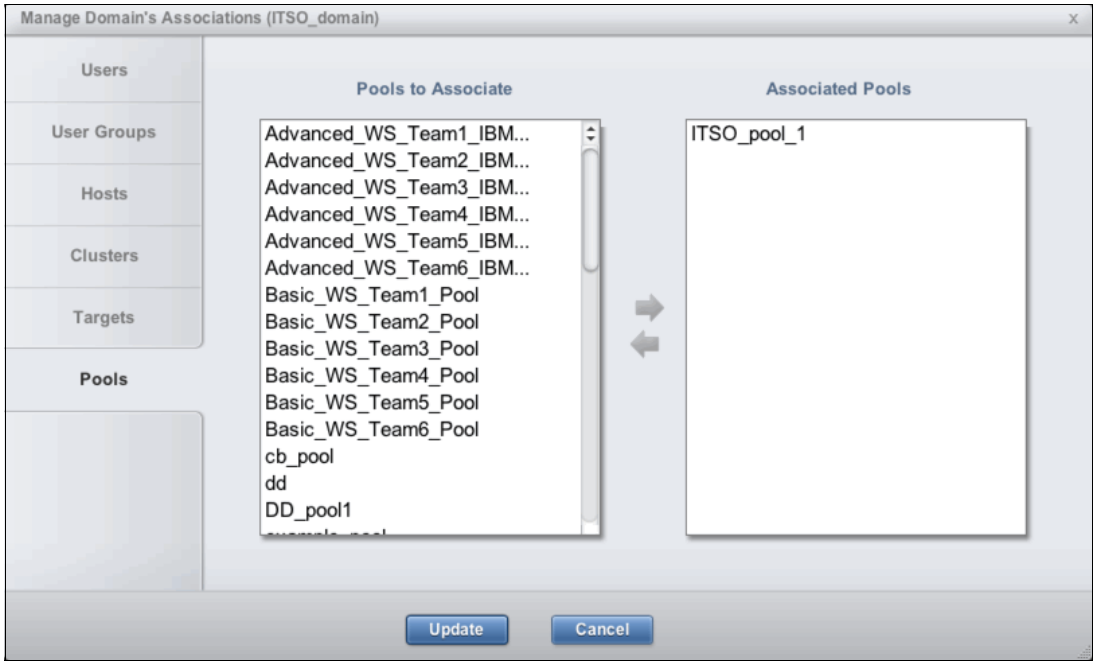


Figure 4-90 Managing domain pool associations in the XIV GUI

Important:

- Associating a pool with a domain might automatically trigger the adjustment of the domain size to accommodate the size of the pool.
- Associating a pool that has pre-existing resource associations (hosts, clusters or targets) with a domain automatically associates those resources with the domain.
- Associating a pool with a domain when it has a pre-existing association with another domain removes the association with the other domain.

For more information about storage pools in XIV, see 2.6, "Storage pool concepts" on page 35.

Managing domain host associations

To manage host associations with a domain, right-click the domain in the Domains view and select **Manage Associations** → **Hosts**, as shown in Figure 4-91 on page 186.

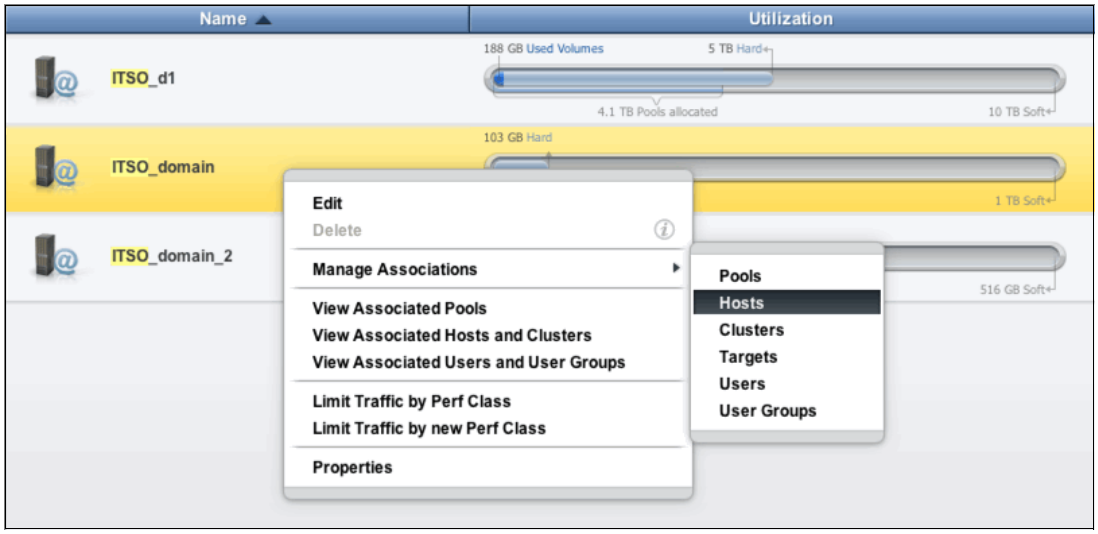


Figure 4-91 Accessing the Manage Domain's Associations screen Hosts tab in the XIV GUI

In the Manage Domain's Associations screen that appears (Figure 4-92), select the hosts and use the arrows to associate or remove the hosts from the domain. Click **Update** to apply the changes.

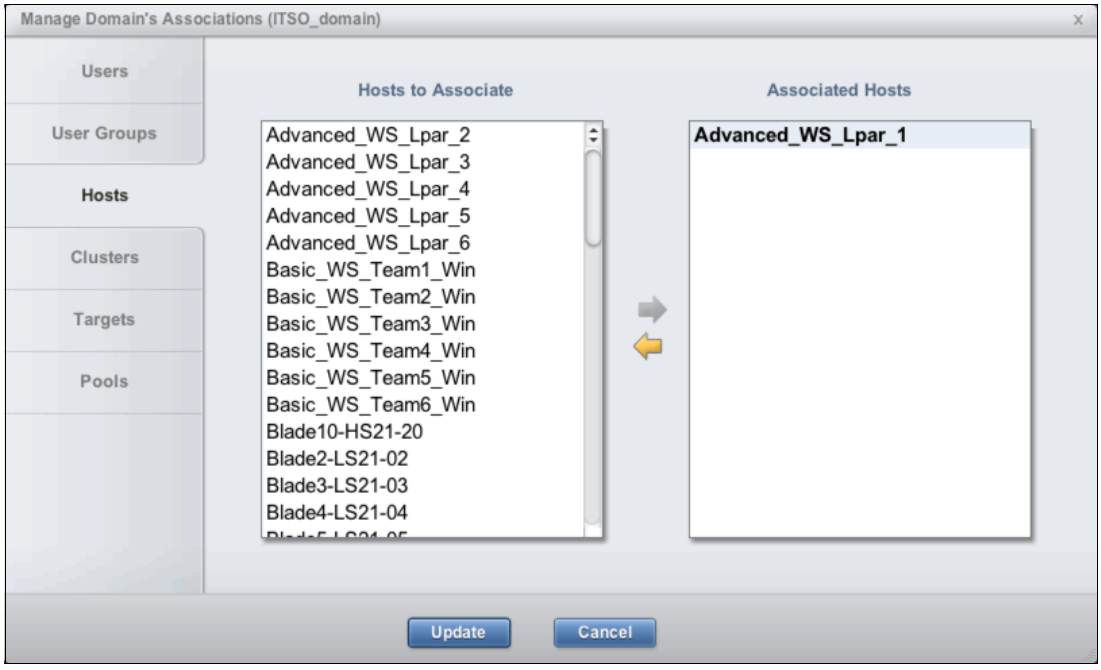


Figure 4-92 Managing domain hosts associations in the XIV GUI

For more information about hosts in XIV, see 4.6, "Host definition and mapping" on page 164.

Managing domain cluster associations

To manage cluster associations with a domain, right-click the domain in the Domains view and select **Manage Associations** → **Clusters**, as shown in Figure 4-93.

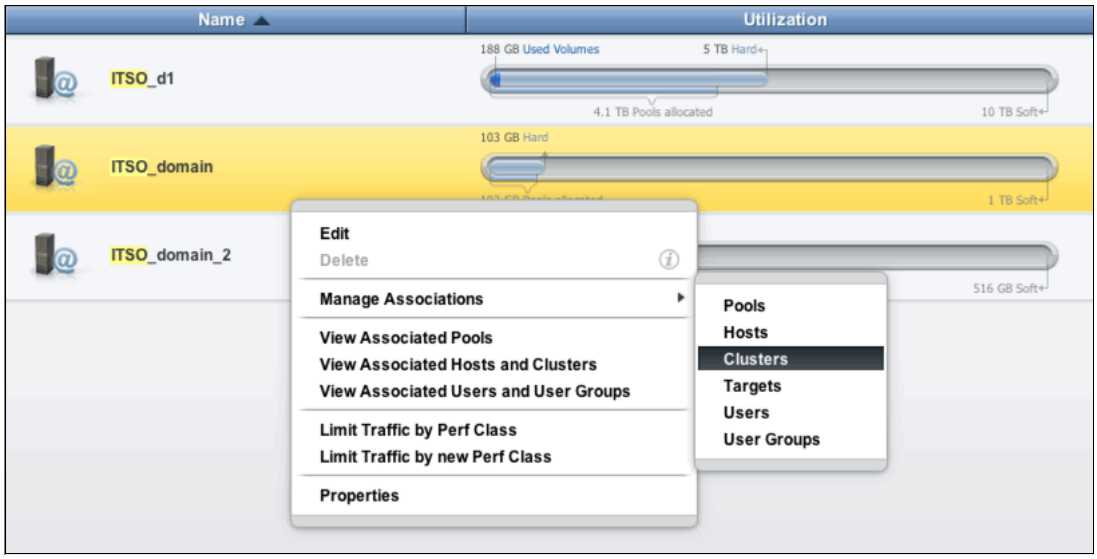


Figure 4-93 Accessing the Manage Domain's Associations screen Clusters tab in the XIV GUI

In the Manage Domain's Associations screen that appears (Figure 4-94), select the clusters and use the arrows to associate or remove the clusters from the domain. Click **Update** to apply the changes.

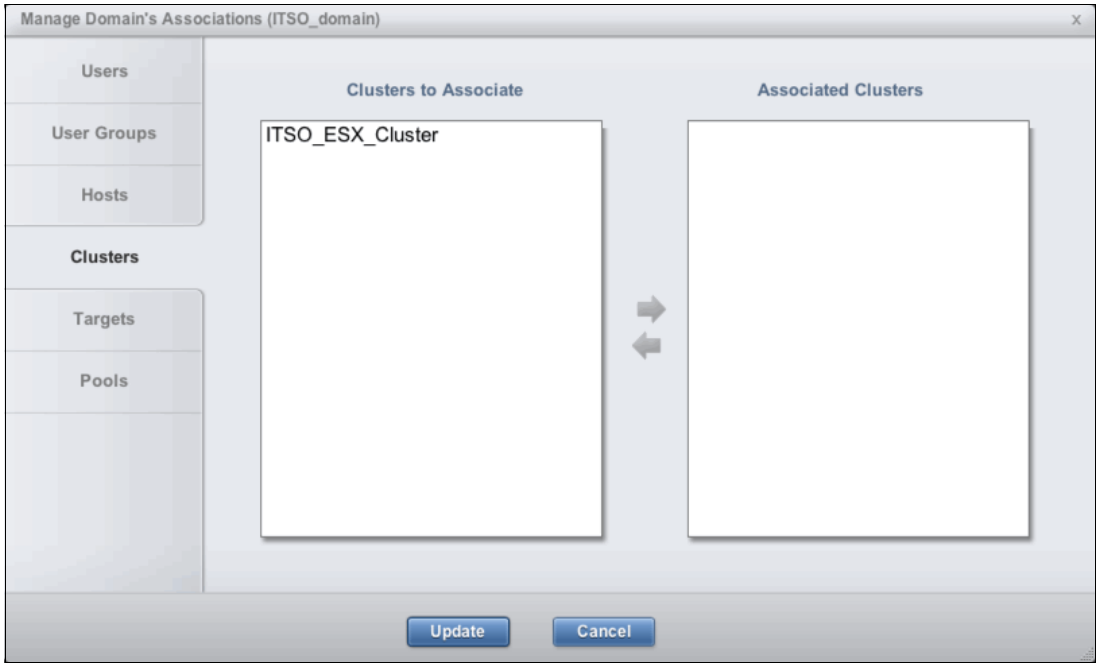


Figure 4-94 Managing domain cluster associations in the XIV GUI

Note: Associating a cluster with a domain automatically associates the hosts contained in the cluster with the domain.

For more information about host clusters in XIV, see 4.6, “Host definition and mapping” on page 164.

Managing domain target associations

To manage remote mirror target associations with a domain, right-click the domain in the Domains view and select **Manage Associations** → **Targets**, as shown in Figure 4-95.

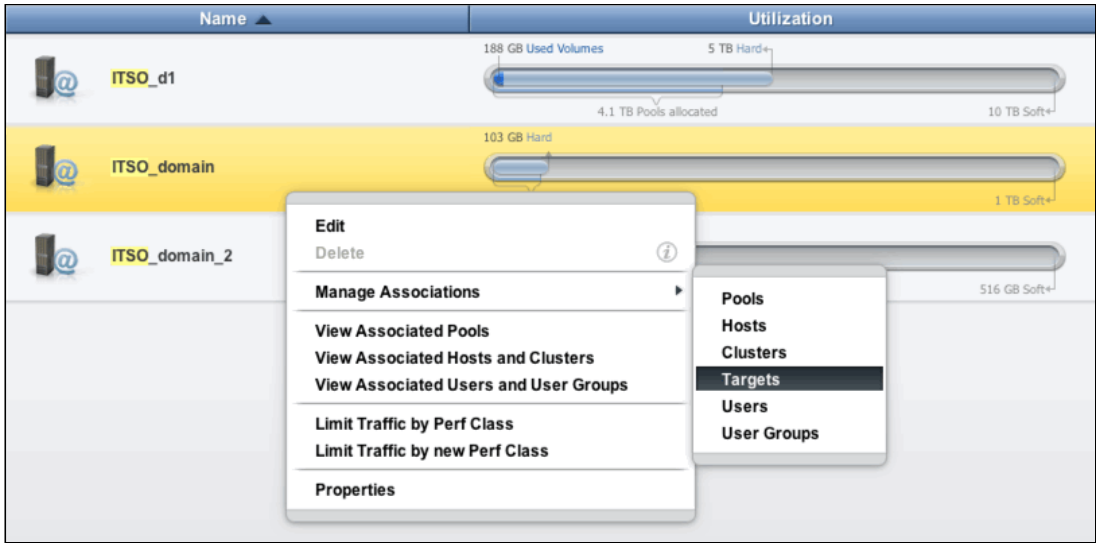


Figure 4-95 Accessing the Manage Domain’s Associations screen Targets tab in the XIV GUI

In the Manage Domain’s Associations screen that appears (Figure 4-96), select the targets and use the arrows to associate or remove the targets from the domain. Click **Update** to apply the changes.

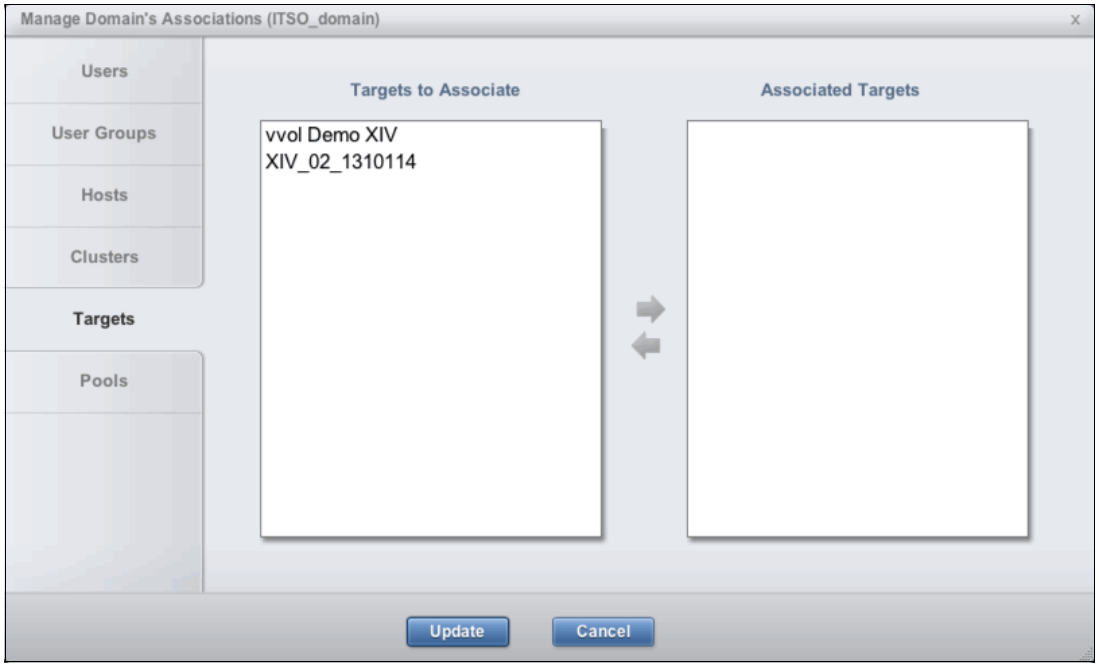


Figure 4-96 Managing domain target associations in the XIV GUI

For more information about remote mirroring and remote mirror targets in XIV, see *IBM XIV Storage System Business Continuity Functions*, SG24-7759.

Managing domain user associations

To manage local user associations with a domain, right-click the domain in the Domains view and select **Manage Associations** → **Users**, as shown in Figure 4-97.

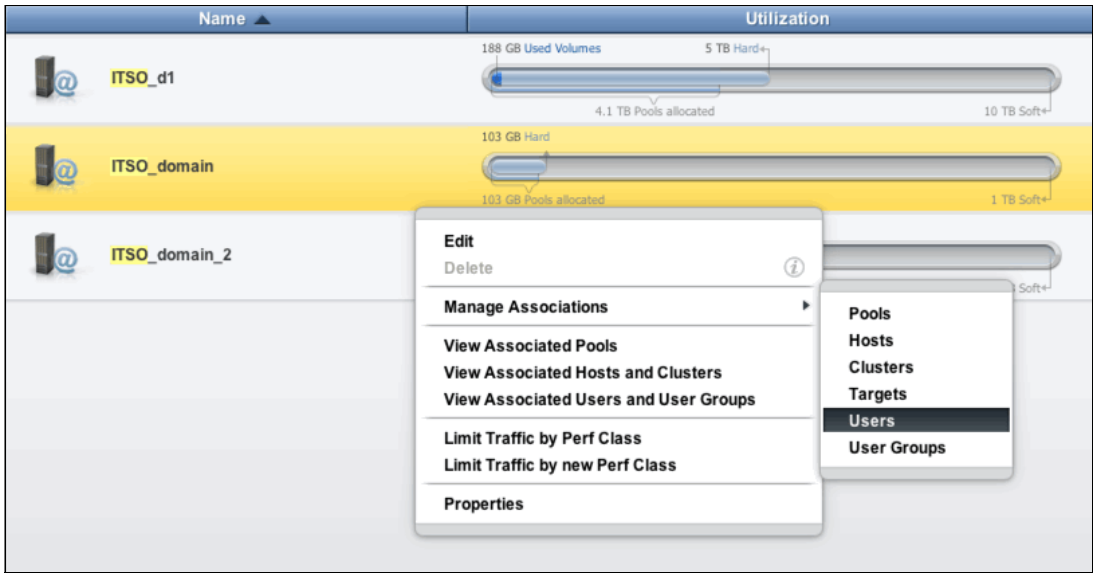


Figure 4-97 Accessing the Manage Domain's Associations screen User tab in the XIV GUI

In the Manage Domain's Associations screen that appears (Figure 4-98), select the users and use the arrows to associate or remove the users from the domain. Click **Update** to apply the changes.

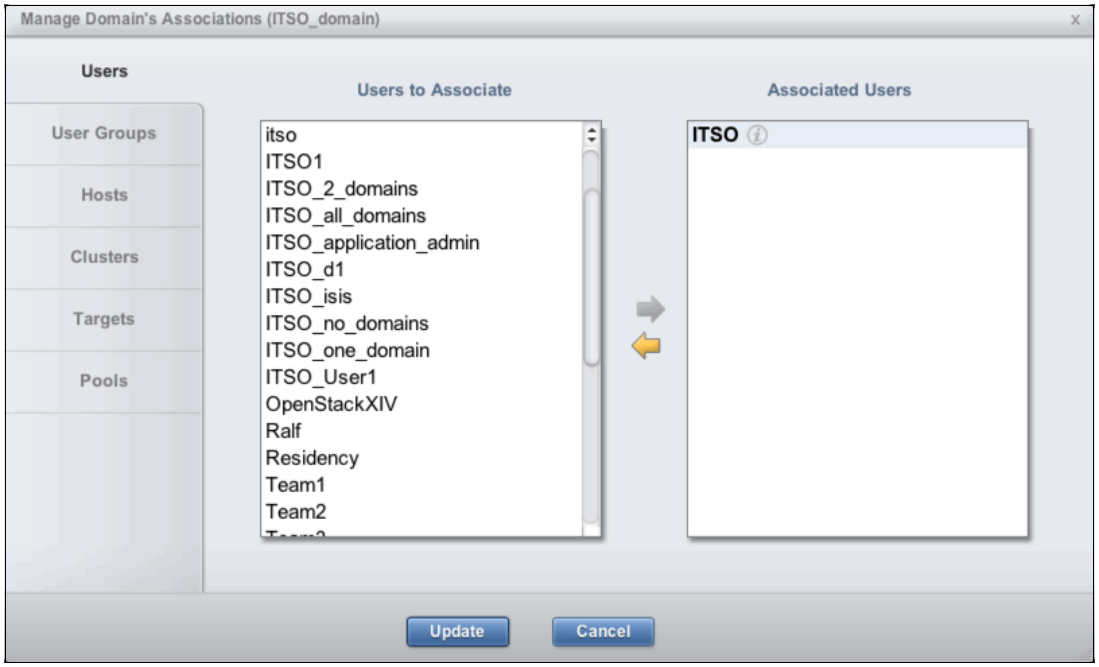


Figure 4-98 Managing domain local user associations in the XIV GUI

Important: Associating a *global* user (a user with no existing domain associations) with a domain will remove the global privileges from that user. Therefore, that user will then have access only to the domains that you associate with that user, explicity. Furthermore, for a local user that is associated with one or more domains, removal of that local user from its last domain association deletes that user from the system. In other words, you cannot convert a *domain* user to a global user.

For more information about XIV user roles, see “User roles” on page 221.

Managing domain user group associations

To manage local user group associations with a domain, right-click the domain in the Domains view and select **Manage Associations** → **User Groups**, as shown in Figure 4-99.

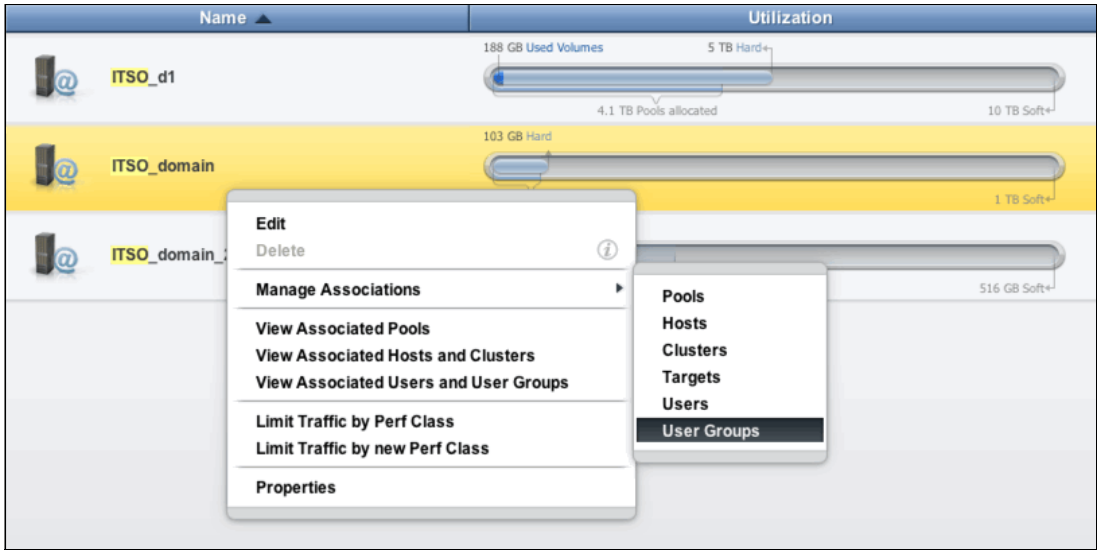


Figure 4-99 Accessing the Manage Domain’s Associations screen User Groups tab in the XIV GUI

In the Manage Domain’s Associations screen that appears (Figure 4-100 on page 191), select the user groups and use the arrows to associate or remove the user groups from the domain. Click **Update** to apply the changes.

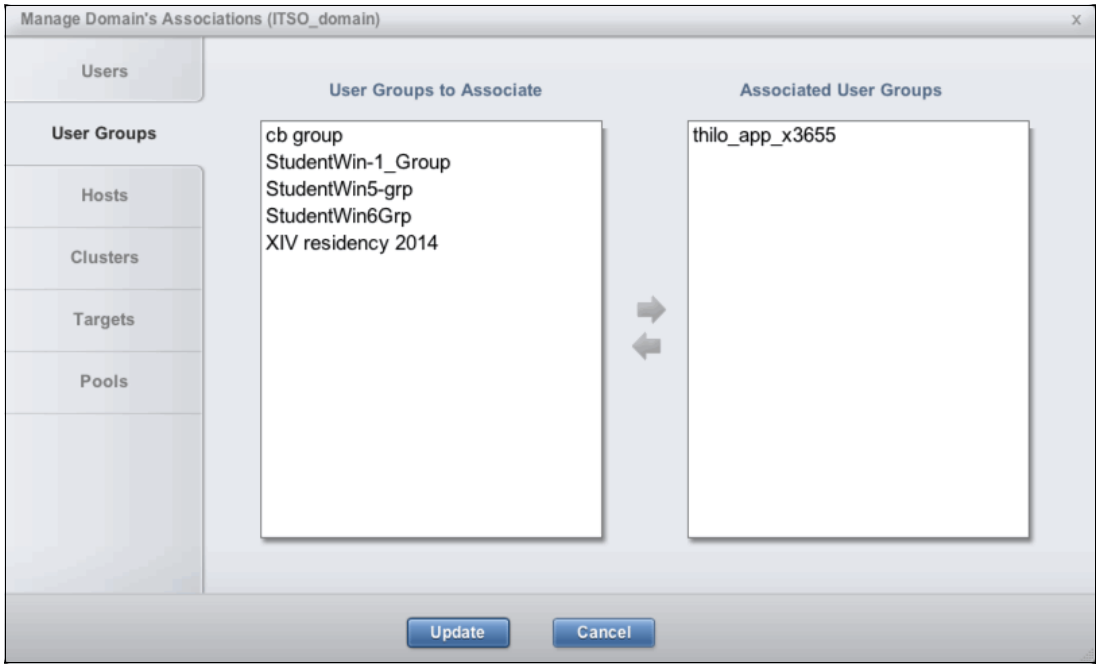


Figure 4-100 Managing domain local user group associations in the XIV GUI

For more information about XIV user groups, see “User groups” on page 224.

Viewing domain associations

You can view the following domain resource associations in the XIV GUI:

- Pools
- Hosts and clusters
- Users and user groups

To view any of these resource associations, right-click the domain in the Domains view and select the appropriate menu item, as shown in Figure 4-101.

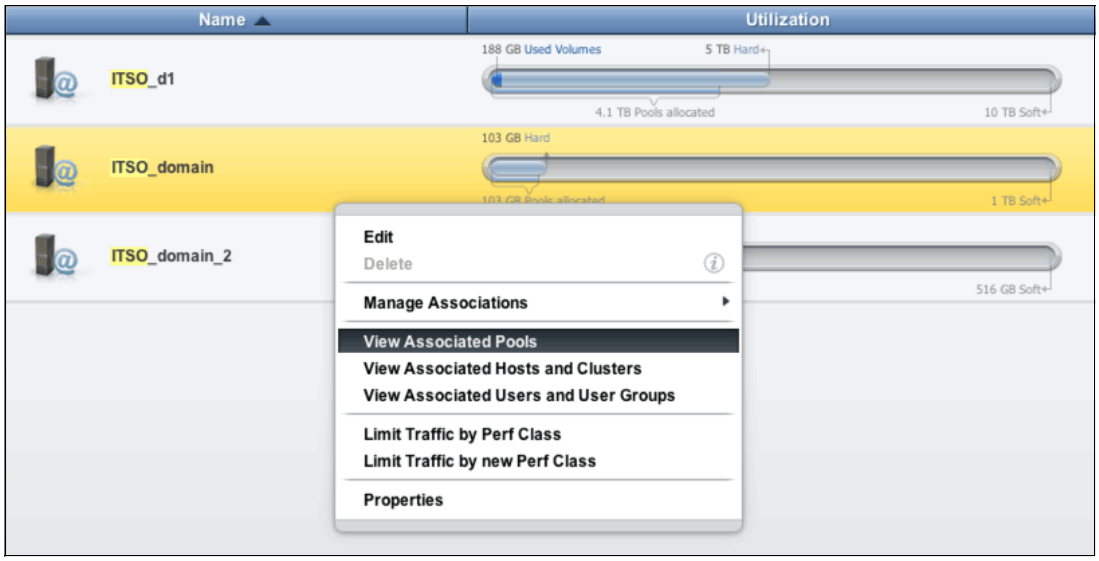


Figure 4-101 Viewing domain resource associations in the XIV GUI

The related view in the XIV GUI will display only those resources that are associated with the domain. As an example, Figure 4-102 shows the Pools view for a single domain called ITSO_domain.

Figure 4-102 Viewing the pool resource associations for a single domain

Deleting a domain

To delete a domain, right-click the domain in the Domains view and select **Delete**, as shown in Figure 4-103.

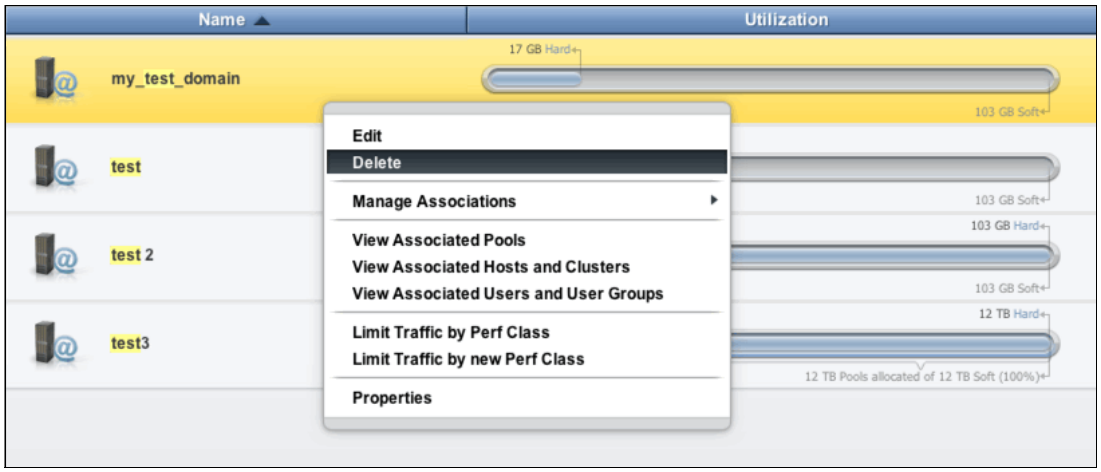


Figure 4-103 Accessing the Delete Domain dialog window in the XIV GUI

To delete the domain, click **OK** in the dialog window that appears, as shown in Figure 4-104.

Figure 4-104 Deleting a domain in the XIV GUI

Note: You cannot delete a domain that contains pools or users. To delete a domain that contains pools or users, first remove the associations to those pools or users.

4.8.3 Limiting domains by performance class (QoS)

In multi-tenant storage environments, to ensure adequate service levels for all tenants, a common requirement is to be able to limit the amount of bandwidth or IOPS that can be consumed by any particular tenant. The XIV Storage System supports the implementation of performance classes (QoS). XIV software Version 11.5 extends the flexibility of the QoS feature, allowing for the application of perfomance classes to domains.

When implementing QoS along with multitenancy, you have the of flexibility for these options:

- Associate a domain with an existing performance class
- Associate a domain with a new performance class

Associating a domain with an existing performance class

To associate a domain with an existing performance class, complete these steps:

1. Right-click the domain in the Domains view, and select **Limit Traffic by Perf Class**, as shown in Figure 4-105.

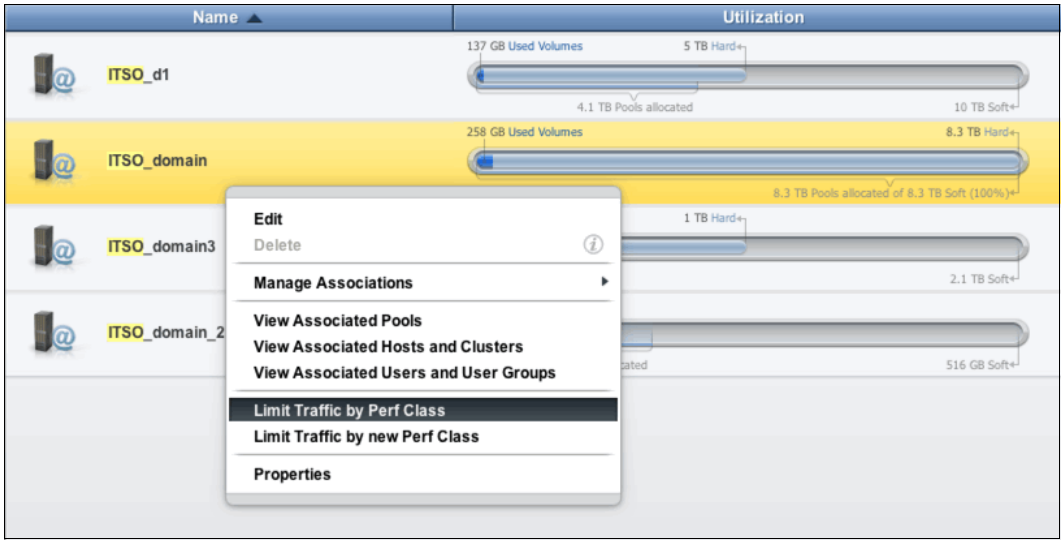


Figure 4-105 Limiting a domain by an existing performance class

2. In the dialog window, select the existing performance class from the drop-down menu and click **Limit**, as shown in Figure 4-106.

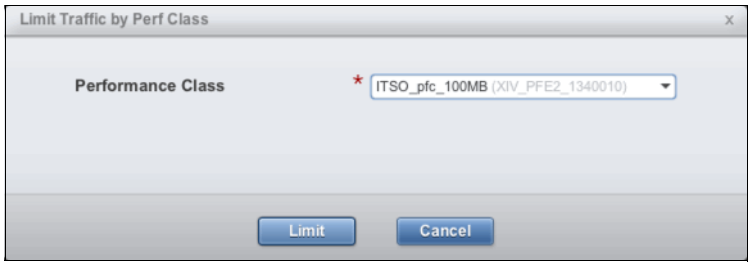


Figure 4-106 Selecting an existing performance class to associate with a domain

3. The GUI opens the QoS Performance class view, where you can confirm that the domain is now associated with the chosen performance class, as shown in Figure 4-107 on page 194.

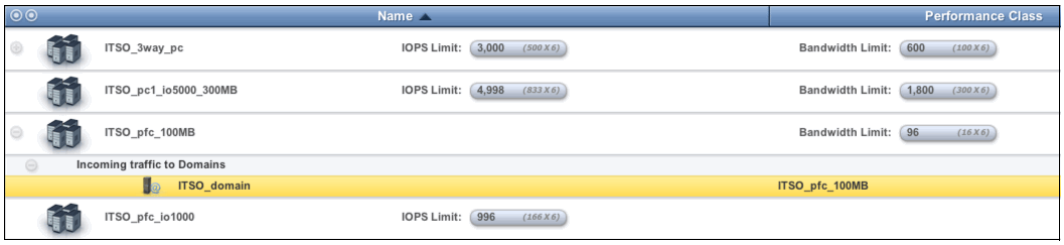


Figure 4-107 A domain successfully associated with an existing performance class

Associating a domain with a new performance class

To associate a domain with a new performance class, follow these steps:

1. Right-click the domain in the Domains view and select **Limit Traffic by new Perf Class**, as shown in Figure 4-108.

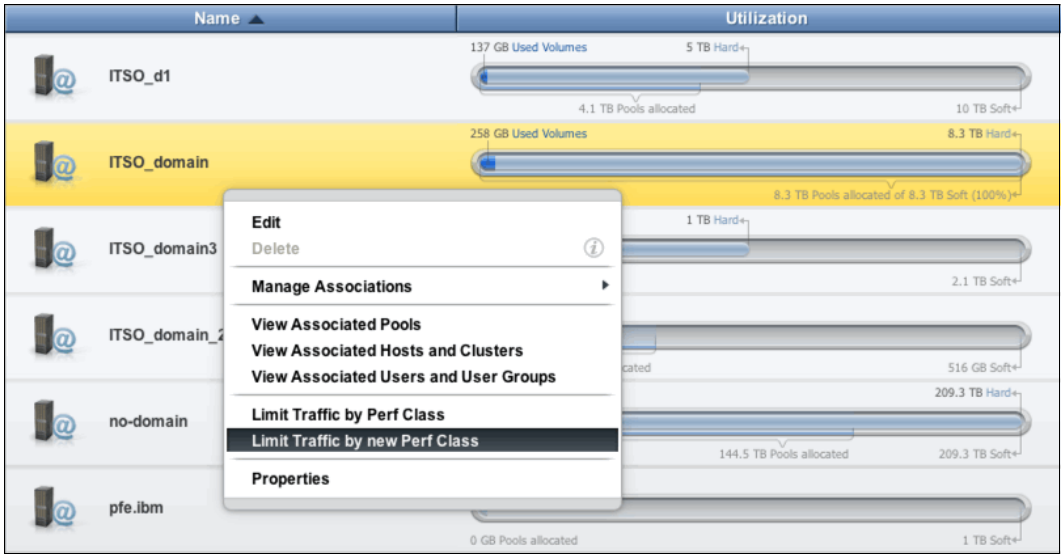


Figure 4-108 Limiting a domain by a new performance class

2. In the dialog window, enter a Name for the new performance class, a value for the I/O Limit or Bandwidth Limit, and click **Add**, as shown in Figure 4-109.

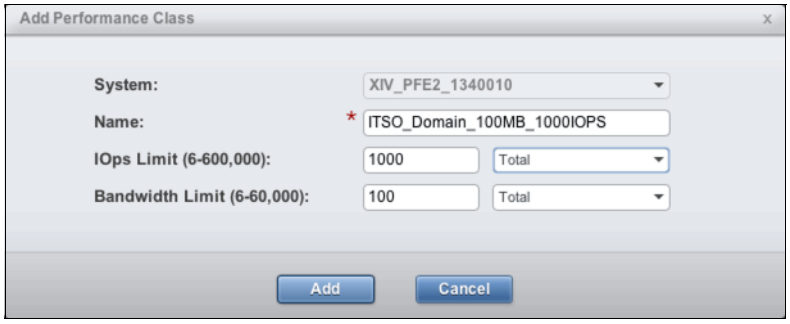


Figure 4-109 Associating a domain with a new performance class

3. The GUI opens the QoS Performance class view, where you can confirm that the domain is now associated with the new performance class, as shown in Figure 4-110 on page 195.

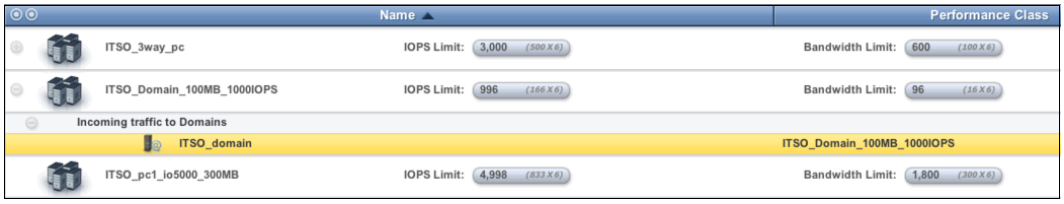


Figure 4-110 A domain successfully associated with a new performance class

4.8.4 Domains and LDAP authentication

As mentioned in 1.6, “XIV Storage System Software” on page 9, the XIV Storage System supports the use of LDAP for user authentication to the system. For more information about configuring the XIV Storage System for LDAP authentication see 5.7, “LDAP-based authentication” on page 244. The use of LDAP authentication along with the new multitenancy feature introduced in Version 11.5 is supported.

With LDAP authentication enabled, it is simple to configure LDAP for use in authentication of users to XIV domains. Figure 4-111 shows the Role Mapping tab of the LDAP parameters window of an XIV Storage System. Notice the value of the highlighted section of the Storage Admin Role parameter. In our example, the highlighted value is *XIV_PFE2_Admins*.

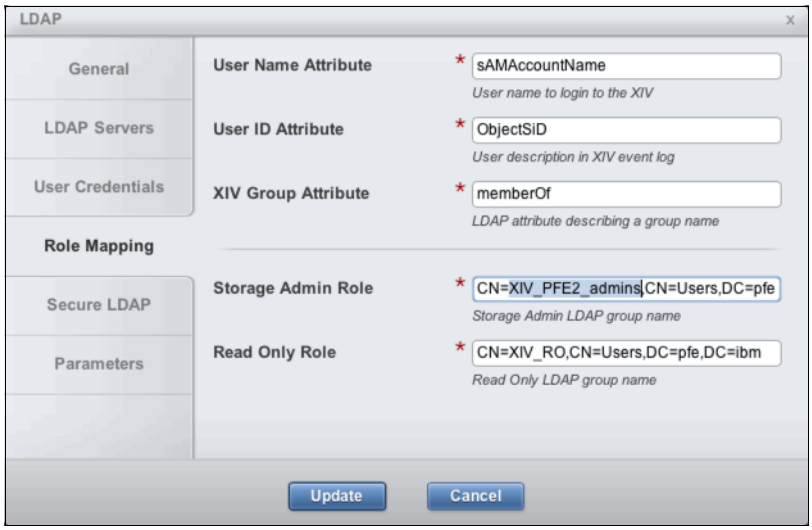


Figure 4-111 The Role Mapping tab in the XIV LDAP parameters window

Without multitenancy, you would create a group in your directory instance name that matches this value, as shown in Figure 4-112 on page 196.

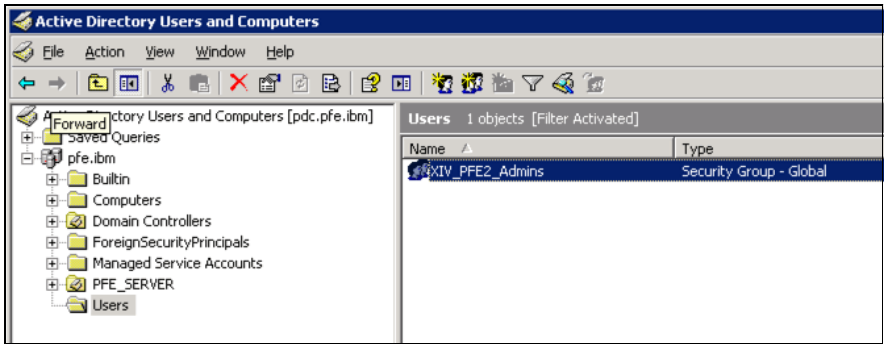


Figure 4-112 An Active Directory group name that matches the XIV Storage Admin Role mapping

To configure a directory group for use with an XIV domain, name the group by using this format:

`role_mapping@LDAP_domain_id`

Where *role_mapping* is the value highlighted in Figure 4-111 on page 195 and *LDAP_domain_id* is the value of the LDAP domain ID defined in the XIV domain properties, as described in “The name, the size, and separated segments are labeled appropriately.” on page 177. Figure 4-113 shows a directory group named using this format.

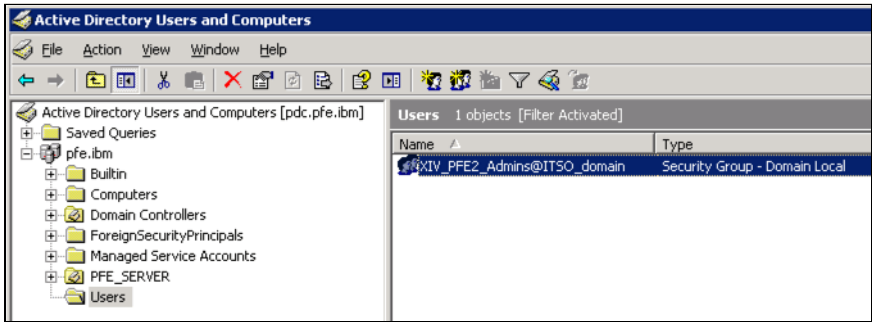


Figure 4-113 An Active Directory group named for use with a single XIV domain

Members of this group in the directory will be granted storage administrator rights in the corresponding domain when they authenticate with the XIV system.

Further, you can configure a directory group for use with multiple XIV domains by appending the additional domains to the end of the name, separating each domain with a / character, as shown in Figure 4-114.

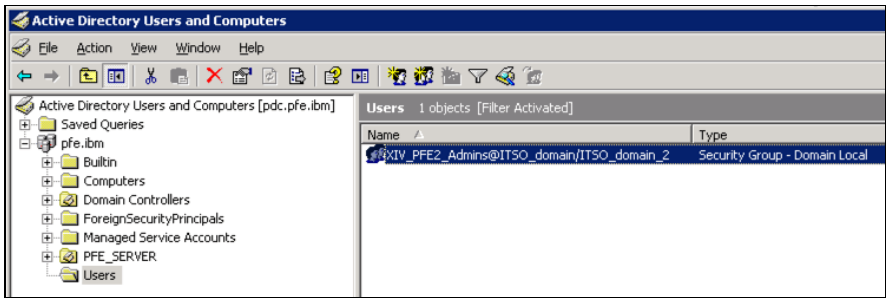


Figure 4-114 An Active Directory group named for use with multiple XIV domains

Note: With XIV multitenancy and LDAP authentication, user membership in multiple groups that are mapped to different roles is not supported. In such a case, the user will not be allowed to authenticate to the XIV system.

Also, remember that domain user associations can be managed only by the Global Storage Administrator through the XIV GUI when local authentication is in use. When LDAP authentication is in use (enabled), domain user associations must be managed directly within the chosen LDAP solution. In addition, with LDAP enabled, a global administrator can associate user groups to a domain but cannot disassociate user groups from a domain with LDAP active.

4.8.5 Managing domains with the XIV command-line interface, XCLI

All of the operations explained in 4.8.2, “Managing domains with the XIV Storage Management GUI” on page 175, can also be performed through the command-line interface. To get a list of all the domain-related commands, enter the following command in an XCLI Session:

```
help search="domain"
```

Important: The commands shown in this section are based on the assumption that you started an XCLI Session on the selected system, as described in “XCLI session features” on page 138. Replace the example parameter values, as appropriate, to perform the commands.

Example 4-17 shows the output of the command.

Example 4-17 All the domain-related commands available in the XCLI

```
XIV_PFE2_1340010>>help search="domain"
Category        Name                        Description
accesscontrol   domain_list_users           Lists users associated with domain(s)
accesscontrol   domain_list_objects         Lists objects attached to domain(s)
accesscontrol   domain_attach_object        Associate Object to a Domain.
accesscontrol   domain_detach_object        Disassociate Object from a Domain.
accesscontrol   domain_policy_get           Get domain-related policies.
accesscontrol   domain_policy_set           Set domain-related policies.
accesscontrol   domain_list                 Lists all Domains or the specified one.
accesscontrol   domain_global_list          Lists the Global Domain.
accesscontrol   domain_create               Creates a Domain.
accesscontrol   domain_update               Updates a Domain.
accesscontrol   domain_delete               Deletes a Domain.
accesscontrol   domain_rename               Renames a Domain.
accesscontrol   domain_add_pool             Associate Pool to a Domain.
accesscontrol   domain_remove_pool          Disassociate Pool from a Domain.
accesscontrol   domain_move_pool            Move Pool from One Domain to Another.
accesscontrol   domain_move_gp              Move Pool from One Domain to Another.
accesscontrol   domain_manage               Move Pool from One Domain to Another.
accesscontrol   domain_add_user             Associate User to a Domain.
accesscontrol   domain_remove_user          Disassociate User to a Domain.
system          perf_class_add_domain       Adds a Domain to a Performance Class.
system          perf_class_remove_domain    Removes a Domain from its Performance Class.
```

To list the existing domains in a system, run the following command:

```
domain_list
```

For a list of users associated with a domain, run the **domain_list_users** command, substituting the appropriate value for the domain parameter, as show in Example 4-18:

Example 4-18 Listing the users associated with a domain with the XCLI

```
XIV_PFE2_1340010>>domain_list_users domain="ITSO_domain"
Domain        User   Category
ITSO_domain   ITSO   storageadmin
```

For a list of the objects associated with a domain, run the **domain_list_objects** command, substituting the appropriate value for the domain parameter, as shown in Example 4-19:

Example 4-19 Listing the objects associated with a domain with the XCLI

```
XIV_PFE2_1340010>>domain_list_objects domain="ITSO_domain"
Domain        Type         Object
ITSO_domain   host         ITSO_x3550-m3-02
ITSO_domain   host         ITSO_x3550-m3-80
ITSO_domain   cluster      ITSO_ESX_Cluster
ITSO_domain   target       XIV_02_1310114
ITSO_domain   schedule     never
ITSO_domain   schedule     min_interval
ITSO_domain   user_group   thilo_app_x3655
```

To create a domain, use the **domain_create** command. Substitute the appropriate value for the domain parameter, as shown in Example 4-20:

Example 4-20 Creating a domain with the XCLI

```
XIV_PFE2_1340010>>domain_create domain="ITSO_domain"
Command executed successfully.
```

Note: When creating a domain from the CLI by using the **domain_create** command, the only required parameter is `domain`, which specifies the name of the domain. Optionally, you can specify the following parameters:

- `hard_capacity` (If omitted, defaults to 0)
- `soft_capacity` (If omitted, defaults to 0)
- `max_pools` (If omitted, defaults to 0)
- `max_volumes` (If omitted, defaults to 0)
- `max_cgs` (If omitted, defaults to 0)
- `max_mirrors` (If omitted, defaults to 0)
- `max_dms` (If omitted, defaults to 0)
- `perf_class` (If omitted, defaults to undefined)
- `ldap_id` (If omitted, defaults to the value of the `domain` parameter)
- `allow_ssd_caching` (If omitted, defaults to `yes` when SSD caching is enabled on the system, but if omitted when SSD caching is disabled on the XIV system, defaults to `no`)

To update a domain, use the **domain_update** command, substituting the appropriate value for the domain parameter, and specifying the parameter values, as shown in Example 4-21:

Example 4-21 Updating a domain with the XCLI

```
XIV_PFE2_1340010>>domain_update domain="ITSO_domain" hard_capacity=1024 soft_capacity=2048
Command executed successfully.
```

4.9 Flash cache

Flash cache is an optional feature introduced and supported by the XIV system software Version 11.1.0 on an XIV Gen3 system. Flash cache adds up to 12 TB of management-free caching power to a fully configured (15-module) storage system.

The flash cache automatically adjusts to the workload, providing a performance boost for small I/O random reads. No additional configuration or tuning is necessary by the storage administrator. However, XIV Storage Manager GUI Version 4 offers enhancements to support flash cache, allowing the storage administrator to perform several optional actions, for example:

- Check the flash cache status in each module.
- Enable or disable flash cache at system level.
- Enable or disable flash cache at volume level.

The remainder of this section shows the essential actions that an administrator can perform through the GUI and the XCLI to manage the flash cache feature. For a complete description of all actions and commands that relate to the flash cache, see the Redpaper publication titled *Solid-State Drive Caching in the IBM XIV Storage System*, REDP-4842.

4.9.1 Managing flash cache with the XIV Storage Management GUI

You can check the health and status of SSDs used for the flash cache in the main system view. Moving the mouse cursor over a module displays a pop-up panel showing the temperature and status of major module components. When SSD disks are present, the SSD status is displayed at the bottom of the pop-up panel. If the SSD is operational, a green `OK` status is displayed as illustrated in Figure 4-115.

Figure 4-115 Module status view with SSD

Clicking a module number opens a full perspective view of the module and its components, as shown in Figure 4-116 on page 200.

From that view, you can also check the SSD status. If an SSD is installed, hover your mouse cursor over the SSD in the module view:

- SSD is highlighted in green if the status is OK and phased in.
- SSD is highlighted in yellow if it is phased out.
- SSD is highlighted in red if it failed.

Figure 4-116 SSD status

Tip: Flash cache can be dynamically enabled and disabled at any time.

Setting flash cache at the system level

By default, flash cache is enabled for all volumes defined in the XIV system.

To enable or disable flash cache for the entire XIV system, select **Settings** from the toolbar in the main System view, as indicated in Figure 4-117.

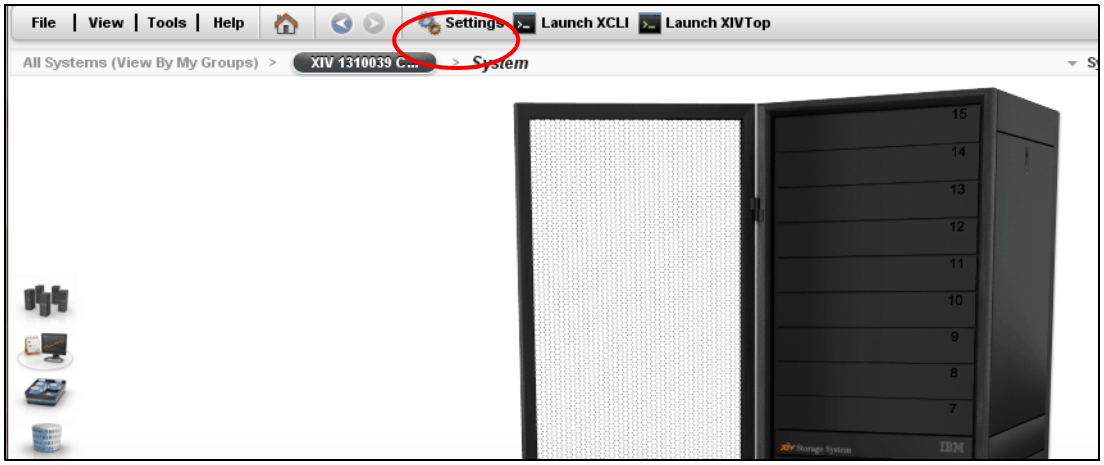


Figure 4-117 Opening system settings

The system Settings panel is displayed. Select the **Parameters** tab. The Global SSD (Flash) Caching default can be set to Enabled or Disabled, as shown in Figure 4-118.

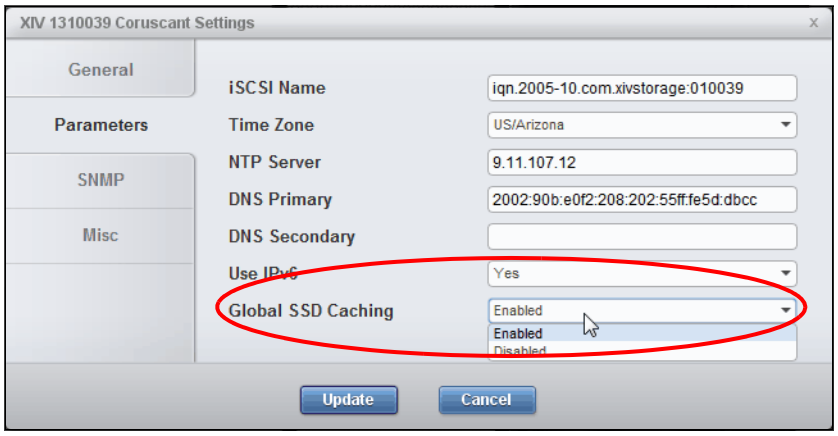


Figure 4-118 System settings for Global SSD Caching

By default, a newly created volume follows the System Level Settings.

System Level Settings:

- If flash cache is disabled from System Level Settings, it is disabled, by default, for every new volume defined.
- If flash cache is enabled from System Level Settings, it is enabled, by default, for every new volume defined.

Setting flash cache at a volume level

The Volumes and Snapshots view in the XIV Management GUI includes a new field that indicates the flash cache status for the volumes that are defined in the system.

In the Volumes and Snapshots tabular view, there is a new column labeled SSD. It displays the flash cache status for all the volumes.

By default, the SSD column shown in Figure 4-119 is not visible. To add it, use the instructions in 4.5.1, “Managing volumes with the XIV Storage Management GUI” on page 153.

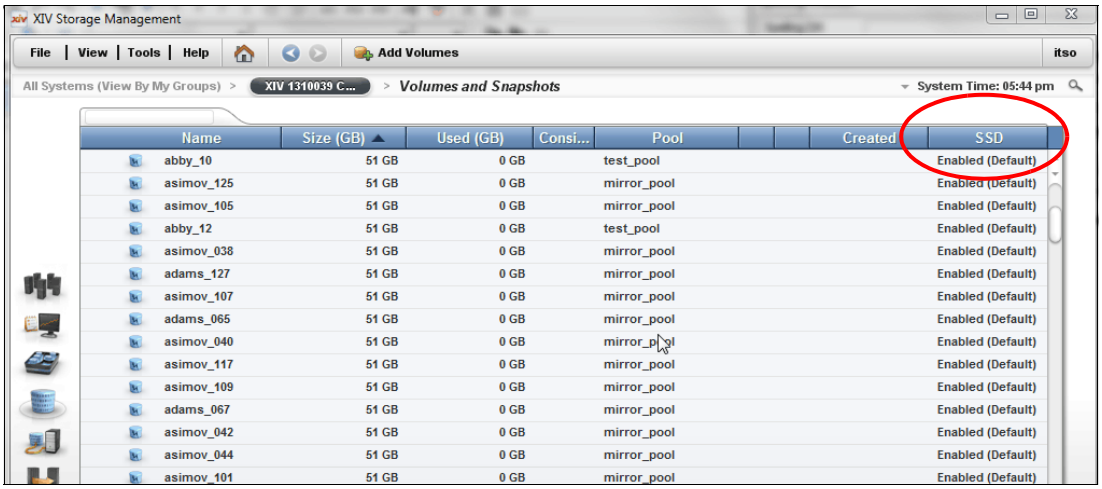

XIV Storage Management
File | View | Tools | Help
Add Volumes
itso
All Systems (View By My Groups) > XIV 1310039 C... > Volumes and Snapshots
System Time: 05:44 pm
Name | Size (GB) | Used (GB) | Consi... | Pool | Created | SSD
abby_10 51 GB 0 GB test_pool Enabled (Default)
asimov_125 51 GB 0 GB mirror_pool Enabled (Default)
asimov_105 51 GB 0 GB mirror_pool Enabled (Default)
abby_12 51 GB 0 GB test_pool Enabled (Default)
asimov_038 51 GB 0 GB mirror_pool Enabled (Default)
adams_127 51 GB 0 GB mirror_pool Enabled (Default)
asimov_107 51 GB 0 GB mirror_pool Enabled (Default)
adams_065 51 GB 0 GB mirror_pool Enabled (Default)
asimov_040 51 GB 0 GB mirror_pool Enabled (Default)
asimov_117 51 GB 0 GB mirror_pool Enabled (Default)
asimov_109 51 GB 0 GB mirror_pool Enabled (Default)
adams_067 51 GB 0 GB mirror_pool Enabled (Default)
asimov_042 51 GB 0 GB mirror_pool Enabled (Default)
asimov_044 51 GB 0 GB mirror_pool Enabled (Default)
asimov_101 51 GB 0 GB mirror_pool Enabled (Default)

Figure 4-119 SSD field on Volumes and Snapshots view

By default, any newly created volume inherits the current system-level cache setting.

System Level Settings:

- If flash cache is disabled from System Level Settings, it is disabled, by default, for every new volume defined.
- If flash cache is enabled from System Level Settings, it is enabled, by default, for every new volume defined.

In the Volumes and Snapshots view shown in Figure 4-119, right-click a volume row and select **Change SSD Caching State** from the pop-up menu.

You are presented a dialog box where you can change the status of the SSD Caching (Figure 4-120).

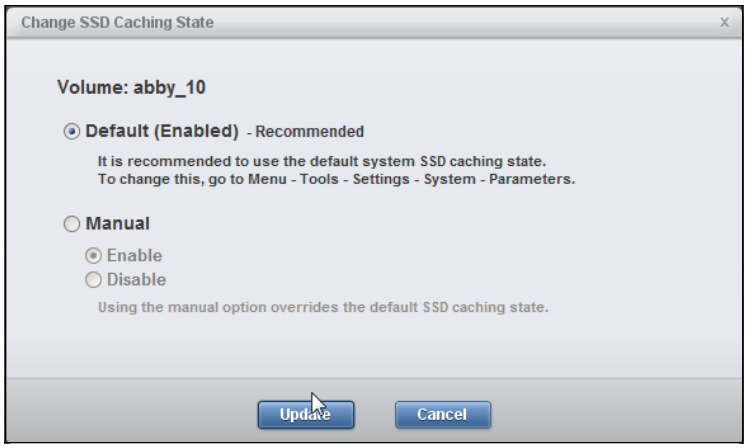

Change SSD Caching State
Volume: abby_10
Default (Enabled) - Recommended
It is recommended to use the default system SSD caching state.
To change this, go to Menu - Tools - Settings - System - Parameters.
Manual
Enable
Disable
Using the manual option overrides the default SSD caching state.
Update
Cancel

Figure 4-120 Change Volume SSD Caching State window

The following information applies to the Change SSD Caching State window:

- If the Default (Enabled) setting is selected, the volume follows the current system settings.
- You can override the Default by selecting **Manual**. Then, the volume no longer follows the current default system settings. In that case, you can select one of these options:
  - **Enable**: Flash cache is enabled for the selected volume.
  - **Disable**: Flash cache is disabled for the selected volume.

Tip: The overall System Status Setting for the SSD Caching is shown in parentheses.

You can select more than one volume in the Volumes and Snapshots view if you need to change the SSD Caching State for a list of volumes.

Flash cache performance statistics

The Performance Statistics view is updated to reflect flash cache.

On GUI 4.0, some new metrics have been added for Read I/O:

- Mem Hit: Metrics for read I/O Main cache
- SSD Hit: Metrics for read I/O Extended SSD cache

The graphic in Figure 4-121 shows a common trend when you enable flash cache. Total Memory Hit (red line) is going up because of flash cache. Meanwhile, Read Miss (blue line) is going down. Flash cache Hit (purple line) is going up side-by-side with Total Memory Hit.

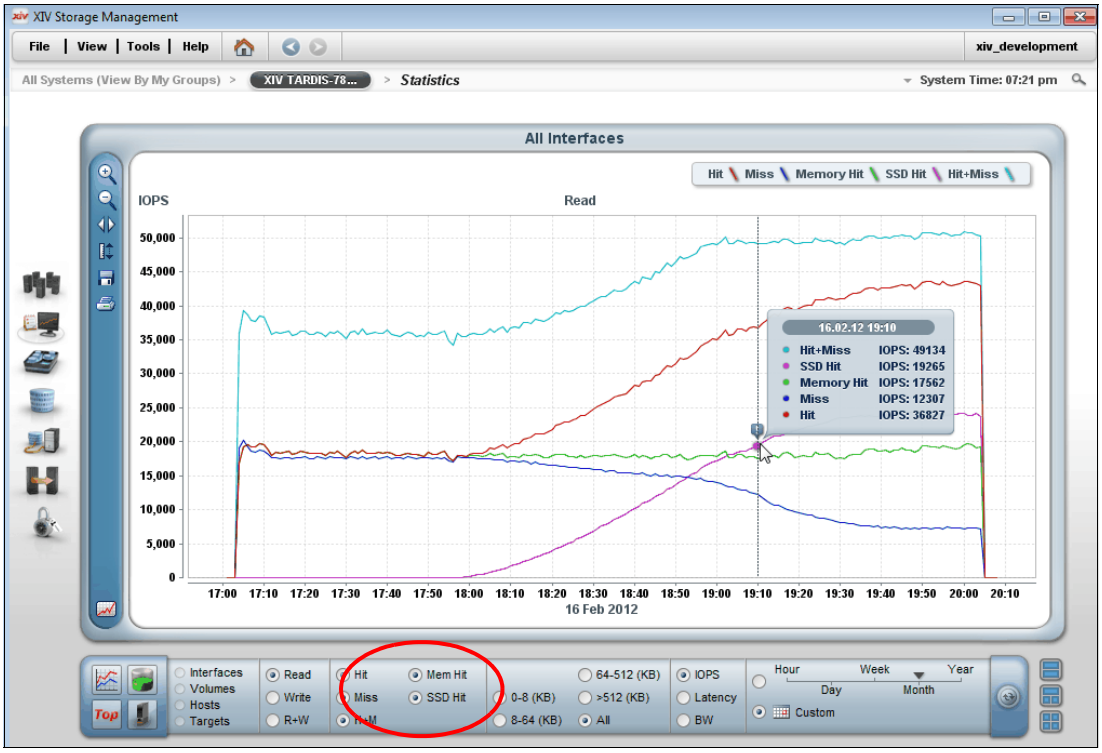


Figure 4-121 Performance Statistics: Mem Hit and SSD Hit

For a description of flash cache performance, see Chapter 6, “Performance” on page 287, and *Solid-State Drive Caching in the IBM XIV Storage System*, REDP-4842.

4.9.2 Managing flash cache with XIV Command Line Interface

New commands are included in the XCLI that matches the possible actions that were previously illustrated for the GUI.

A useful command is **`help search=ssd`**, which displays a list of all commands related to SSDs. This command is illustrated in Example 4-22.

Example 4-22 List of commands related to SSDs

```
XIV 1310039 Coruscant>>help search=ssd
Category   Name                          Description
system     ssd_caching_disable           Disables Flash Caching
system     ssd_caching_enable            Enables SSD Caching
system     ssd_list                      Lists SSDs used as flash cache in the
                                           system
system     vol_default_ssd_caching_get   Gets the Default State of the SSD Caching
system     vol_default_ssd_caching_set   Sets a Default State for SSD Caching
system     vol_ssd_caching_set           Overrides the Default SSD Caching State
                                           for a Volume
```

The **ssd_caching_disable** and **ssd_caching_enable** commands, which are listed in Example 4-22 on page 204, are restricted commands that can be used only with the IBM technician's authority. These commands are used by the IBM technician to bring SSDs online when they are phased into the XIV system at installation time.

Use the **ssd_list** command to get a list of SSDs that are used as flash cache in the system. The output is shown in Example 4-23.

Example 4-23 List of SSDs

```
XIV 1310039 Coruscant>>ssd_list
Component ID   Status   Currently Functioning   Capacity   Target Status   Vendor
Model                  Size   Serial   Firmware   Fru       Group   Temperature
1:SSD:3:1      OK       yes                     512GB                    XIV
MTFDDAA512MAR-1KAAB   488378 0200BB81 MA40       99Y0720           42
1:SSD:2:1      OK       yes                     512GB                    XIV
MTFDDAA512MAR-1KAAB   488378 0200BB78 MA40       99Y0720           40
1:SSD:6:1      OK       yes                     512GB                    XIV
MTFDDAA512MAR-1KAAB   488378 0200BB71 MA40       99Y0720           47
1:SSD:4:1      OK       yes                     512GB                    XIV
MTFDDAA512MAR-1KAAB   488378 0200BB95 MA40       99Y0720           47
1:SSD:1:1      OK       yes                     512GB                    XIV
MTFDDAA512MAR-1KAAB   488378 0200BB7A MA40       99Y0720           39
1:SSD:5:1      OK       yes                     512GB                    XIV
MTFDDAA512MAR-1KAAB   488378 0200BB7D MA40       99Y0720           43
```

System-level flash cache

The **vol_default_ssd_caching_get** command is used to check the default flash cache setting (enabled or disabled), as illustrated in Example 4-24.

Example 4-24 Check SSD Caching state

```
XIV 1310039 Coruscant>>vol_default_ssd_caching_get
Command executed successfully.
default=enabled

XIV 1310039 Coruscant>>vol_default_ssd_caching_get
Command executed successfully.
default=disabled
```

If the default status is *enabled*, flash cache is enabled on all volumes in the system unless the status is manually changed. Otherwise, if default status is *disabled*, flash cache is disabled for all volumes in the system.

Tip: With the flash cache state enabled, you can explicitly disable any volume that you do not want to include in the extended caching.

If you want to change the flash cache state globally for the system, you can issue one of the following commands:

```
vol_default_ssd_caching_set default=enabled
set the SSD Caching enabled

vol_default_ssd_caching_set default=disabled
set the SSD Caching disabled
```

Volume-level flash cache

Use the **vol_ssd_caching_set** command to set the flash cache status for a specific volume, and eventually override the system default setting. You must specify the **vol** and **state** parameters when you issue the command, as illustrated in Example 4-25.

Example 4-25 SSD Caching set for a volume

```
XIV 1310039 Coruscant>>vol_ssd_caching_set vol=Res_Fra_Vol_01 state=enabled
Command executed successfully.

XIV 1310039 Coruscant>>vol_ssd_caching_set vol=Res_Fra_Vol_01 state=disabled
Command executed successfully.
```

You can also use the **vol_list** command with the **-x** flag and the **vol** parameter to display all of the volume properties, which now include two additional values, **ssd_caching** and **use_ssd_caching_default**. For an illustration, see Example 4-26.

Example 4-26 State of a volume

```
XIV 1310039 Coruscant>>vol_list -x vol=Res_Fra_Vol_02
<XCLIRETURN STATUS="SUCCESS" COMMAND_LINE="vol_list -x vol=Res_Fra_Vol_02">
  <OUTPUT>
            <volume id="bec1240012f">
                <creator value="itso"/>
                <creator_category value="storageadmin"/>
                <id value="bec1240012f"/>
                <name value="Res_Fra_Vol_02"/>
                <size value="17"/>
                  ............
                  ............
               <ssd_caching value="disabled"/>
                <use_ssd_caching_default value="yes"/>
            </volume>

  </OUTPUT>
</XCLIRETURN>
```

The value of the **use_ssd_caching_default** parameter indicates whether the volume follows the default system state for flash cache:

- If the value is `yes`, the volume follows the default system state for flash cache.
- If the value is `no`, the volume does not inherit the default setting for flash cache. It means that if the global system setting for the caching is changed, the volume keeps its current **ssd_caching** value.

5

Security

This chapter describes the IBM XIV Storage System security features from various perspectives. It covers the following topics:

- Physical access security
- x509 certificate validation and management
- Configuring IPv6 addresses
- Configuring Internet Protocol Security (IPSec) connectivity
- Native user authentication
- Considerations when using MultiSystem Manager
- LDAP-based authentication
- Defining LDAP on the XIV Storage System
- LDAP-managed user authentication
- Securing LDAP communication with Secure Sockets Layer

Important: An important aspect of XIV security is the support for data-at-rest encryption, which was introduced with XIV Storage Software Version 11.4. The encryption topic is covered in the IBM Redpaper publication titled, *IBM XIV Security with Data-at-Rest Encryption*, REDP-5047.

© Copyright IBM Corp. 2014. All rights reserved.

5.1 Physical access security

When installing an XIV Storage System, you need to apply the same security good practices that you apply to any other business critical IT system. A good reference on storage security can be found at the Storage Networking Industry Association (SNIA) website:

http://www.snia.org/forums/ssif/programs/best_practices

A common risk with storage systems is the retention of volatile caches. The XIV Storage System is perfectly safe in regard to external operations and a loss of external power. If there is a power failure, the internal uninterruptible power supply (UPS) units provide power to the system. The UPS enables the XIV Storage System to gracefully shut down.

However, if someone gains physical access to the equipment, that person might manually shut off components by bypassing the preferred process. In this case, the storage system is likely to lose the contents of its volatile caches, resulting in a data loss and system unavailability. To eliminate or greatly reduce this risk, the XIV Storage System rack can be equipped with lockable doors. The XIV Gen3 rack security kit is available by ordering RPQ 8S1190.

Important: Protect your XIV Storage System by locking the rack doors and monitoring physical access to the equipment.

5.2 x509 certificate validation and management

All communication among the XIV graphical user interface (GUI), XIV Storage System Command-Line Interface (XCLI), Hyper-Scale Manager, and XIV systems uses Secure Sockets Layering (SSL) via x509 certificates.

The use of x509 certificates provides for secure authentication and encryption of all communication between the XIV software components. Previous versions required the use of the default, built-in certificate. Beginning with XIV Storage System Software Version 11.2 and XIV Storage Management software Version 4.1, you have the flexibility to install and use your own x509 certificates in addition to the built-in certificate.

You can choose to use a certificate that is signed by either a trusted certificate authority (CA) vendor or your organization's own private CA server. For more information about creating and using your own CA server, see "Certificate authority setup" on page 460.

Four steps are required to use your own certificate:

1. Generate a certificate signing request (CSR) file from the XIV Storage System.
2. Obtain a signed certificate using the CSR file, either from a CA vendor, or with your own CA server.
3. Install the signed certificate on the XIV Storage System.
4. Instruct the XIV GUI to trust the new signed certificate.

5.2.1 Managing x509 certificates with the XIV Storage Management GUI

This section shows how to prepare and configure the XIV Storage System to use x509 certificates.

Creating a CSR and importing the signed x509 certificate

The following steps demonstrate how to use the XIV GUI to use your own x509 certificate:

1. In the GUI, navigate to the System view of the chosen XIV Storage System and click **Systems** → **System Settings** → **Manage Certificates** to open the Certificates Management panel shown in Figure 5-1.

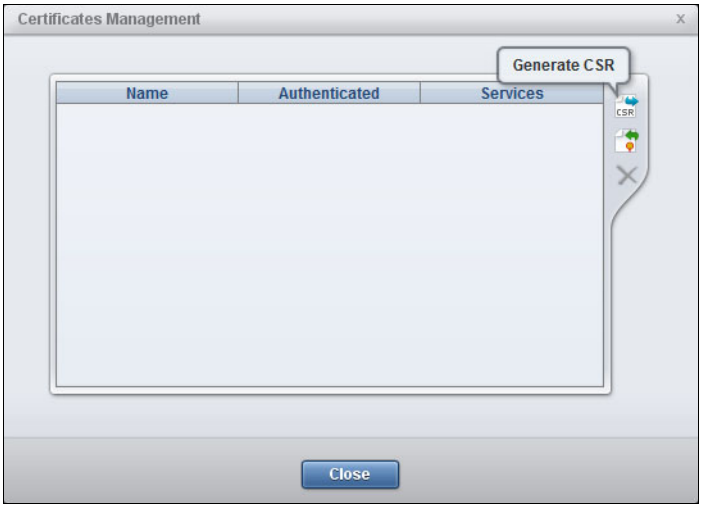


Figure 5-1 Certificates Management panel

2. Click the **Generate CSR** icon in this panel to open the Generate CSR panel, as shown in Figure 5-2.

Figure 5-2 Generate CSR file

Enter a unique value for the name of the certificate in the Name field. In our example, we use the value `itsoCertificate` for the name.

In the Subject field, enter a value for the subject of the certificate. The subject field represents the values that uniquely identify this system, and are commonly called, collectively, the *distinguished name* (*DN*). The acceptable format for the subject field is a string of attribute=value pairs, each preceded by a slash. Spaces are not permitted. In our example, we use the value `/CN=xivhost/O=itso/L=Tucson/ST=AZ/C=US`.

Choose a value for the encryption strength in the Bits field. In our example, we choose 2048-bit encryption.

Tip: The subject field in an x509 certificate uniquely identifies the host that the certificate belongs to. It is recommended that, at a minimum, the following field attributes be included in the subject or your certificate request:

- CN (common name)
- O (organization)
- L (locality)
- ST (state)
- C (country)

Check with your CA vendor or CA server administrator to determine whether they require any additional attributes in the subject of your certificate request.

3. Click **Generate** to generate the CSR file. A file browser window opens, as shown in Figure 5-3, prompting you to save the CSR file to your local workstation. Choose the appropriate location and save the CSR file. You provide this file to your CA to produce a signed certificate.

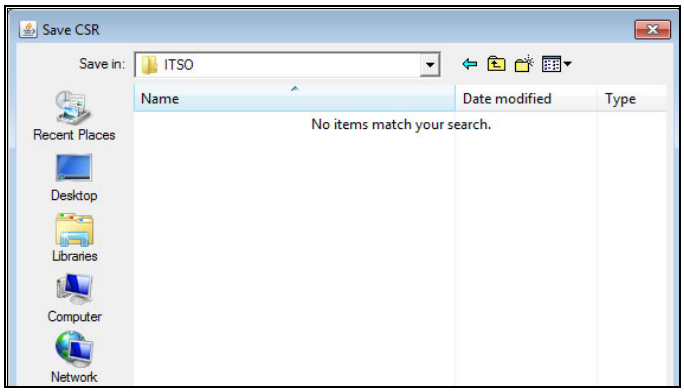


Figure 5-3 Save the CSR file

After saving the CSR file, the Certificates Management panel shows your pending certificate request, as shown in Figure 5-4.

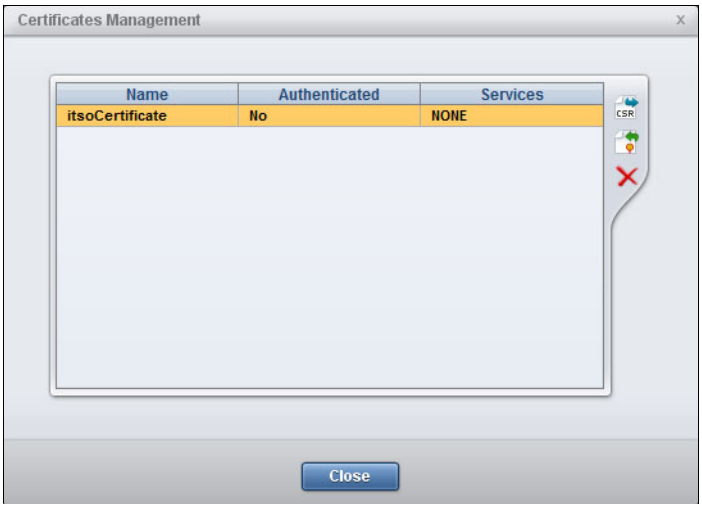


Figure 5-4 Pending CSR

4. Use the CSR file that you just created to obtain a signed x509 certificate, either from a trusted CA vendor, or from your organization's own CA server. For more information about setting up your own CA server and creating your own signed certificate, see "Certificate authority setup" on page 460.

Tip: Depending on what certificate authority server is used to create your signed privacy enhanced mail (PEM) file, the contents of the file might differ slightly. For example, when OpenSSL is used to generate a signed PEM file, the file might contain plain-text metadata details about the CA server and the certificate.

To import the PEM file into an XIV Storage System, it must conform to the x509 standard PEM file format. That is, it should contain only the actual encrypted certificate and the enclosing **-----BEGIN CERTIFICATE-----** and **-----END CERTIFICATE-----** tags, as shown, below:

```
-----BEGIN CERTIFICATE-----

MIIEFTCCAv2gAwIBAgIBJzANBgkqhkiG9w0BAQUFADBLMQswCQYDVQQGEwJVUzELMAkGA1UECBMC
QVoxDDAKBgNVBAoTA0lCTTENMAsGA1UECxMESVRTTzESMBAGA1UEAxMJeGl2Q0Fob3N0MB4XDTEz
MDMyMDIzNTk0MVoXDTE0MDMyMDIzNTk0MVowTDELMAkGA1UEBhMCVVMxCzAJBgNVBAgTAkFaMQ8w
DQYDVQQHEwZUdWNzb24xDTALBgNVBAoTBGl0c28xEDAOBgNVBAMTB3hpdmhvc3QwggEiMA0GCSqG
SIb3DQEBAQUAA4IBDwAwggEKAoIBAQC4hOTzzsk9W+oMfd9+3qwZchx2ZbUjDmbwhW5jyf+9qbb+
at6UeqavD2mGTX/fceLo7ZwFC5M5PL1LMiD+Scz2FWaMH58srwwBo5vUvM/3/P+Du+H1Xb9FwoyF
uUAyIpNkaoMfjL96ToF9NLZ22PTiO48e3Tnk4d/trL1r2kt1fzBf5VChAl79K9aMm+N7PFkjuWJu
vBPSySyCZGhuTLzOERO4xN9zwXHrhohSnBwOZV+kN5NgEVZ6K+s+OtUheksEo/4Mqmhnu3+oOxjH
PYHM7Wu9HrYZU2F+Dm2byrl4ZOL9IcHNd+aCMtraJ6/N6nPiGeFbRS7uUTPmA9VOTf/7AgMBAAGj
ggEBMIH+MAkGA1UdEwQCMAAwEQYJYIZIAYb4QgEBBAQDAgZAMCwGCWCGSAGG+EIBDQQfFh1PcGVu
U1NMIEdlbmVyYXRlZCBDZXJ0aWZpY2F0ZTAdBgNVHQ4EFgQU2imuYrUOFeaVUnsjk9jhajwAMsww
ewYDVR0jBHQwcoAUMsOyREEVImhVvzylyXAUSYsca7ehT6RNMEsxCzAJBgNVBAYTAlVTMQswCQYD
VQQIEwJBWjEMMAoGA1UEChMDSUJNMQ0wCwYDVQQLEwRJVFNPMRIwEAYDVQQDEwl4aXZDQWhvc3SC
CQCq4OFWskg2IDAJBgNVHREEAjAAMAkGA1UdEgQCMAAwDQYJKoZIhvcNAQEFBQADggEBAGqhgptA
I9QB/IuMirBN9QyzUwyigpGcxnwxcSUwRHbC+aqoB4jGSZszdOM3kXctAve1iOhh7t/bJvN5yYHc
/SaojSPjQYyQNHK6FuNVBlDqkLUrRDoX6BtPUBvfQOasZIVQhY4oOObUJg1/GscSuzUkt6m9/sOg
YSw57qDRRD7jHwzxXCzBpNjKOnVcSu/Hy8XUa78z871KCkZKxcc+t2HV7InrBbVDzcBuOrerZuuT
mzLrmzuoNluo7isIUqjhAXf3oSwa+9NVvXFQYTJxiyxIA5MOil3tBnqWQ6PJUKwTbntf7Ra32W5g
zTdA7jDMNSItVfqaM/1bXrk5cOy/Eig=

-----END CERTIFICATE-----
```

If the PEM file that you receive contains any other details, open the file with a text editor, remove the additional details, and save the PEM file before you import it via the XIV GUI.

5. When you have obtained the signed certificate, return to the XIV GUI to import the certificate into your XIV Storage System. Click **Systems** → **System Settings** → **Manage Certificates** to open the Certificates Management panel.

   In the panel, click the **Import Certificate** icon, as shown in Figure 5-5.

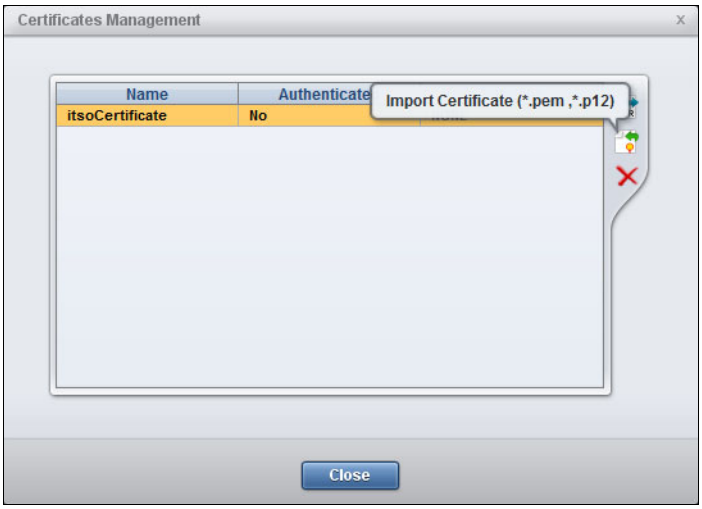


Figure 5-5 Import certificate (PEM file)

6. After the Import Certificate window opens shown in Figure 5-6 opens, click **Browse** to open a file browser window and select the signed certificate file. In our example, the signed certificate file is named `itsoCertificate.pem`.

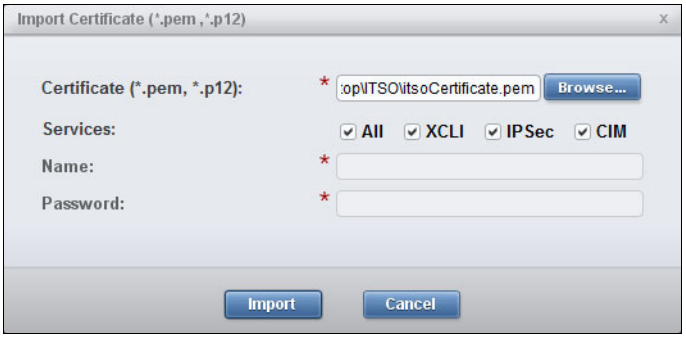


Figure 5-6 Import signed certificate (PEM file)

Click the corresponding check boxes to select the services that you would like to use this certificate for. The following options are available:

All Use this certificate to secure all communications.

XCLI Use this certificate to secure XCLI communication only.

IPSec Use this certificate to secure IPSec traffic (for more information about IPSec, see 5.4, “Configuring Internet Protocol Security connectivity” on page 217).

CIM Use this certificate to secure Common Information Model (CIM) agent communications only.

If the certificate type is PCKS12 (.p12 file extension), enter values for the Name and Password for the certificate.

Click **Import** to complete the import of the signed certificate into the XIV Storage System. The successfully imported certificate is shown in Figure 5-7.

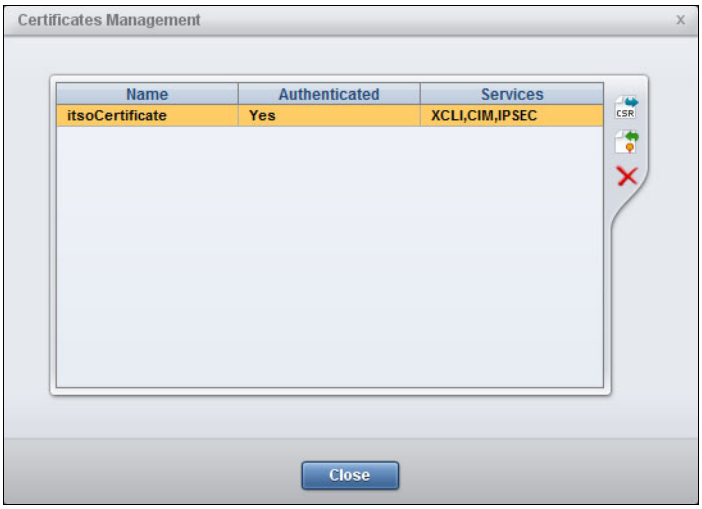


Figure 5-7 Imported certificate

7. Optionally, repeat steps 1 - 6 for any additional IBM XIV systems that you want to secure with your own certificate.

Trusting the signed x509 certificate in the XIV GUI certificate

When you have imported your own certificate into the XIV Storage System, that system will present this certificate to the management tools (GUI, XCLI, and so on) that you use to connect to that system. Because this is a new certificate, the GUI reports a certificate error when connecting to that system, as shown in Figure 5-8.

Figure 5-8 GUI certificate error

Proceed to resolve the certificate error:

1. To resolve this error, you must configure the GUI to trust the new certificate. To do so, right-click the system and choose **Manage Certificate**, as shown in Figure 5-9.

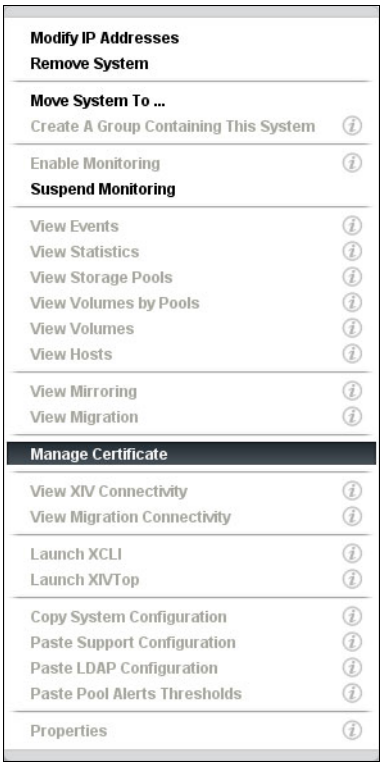


Figure 5-9 Manage Certificate menu option

2. The GUI displays the details of the new certificate, as shown in Figure 5-10. Click **Trust Always** to trust this certificate for all future connections.

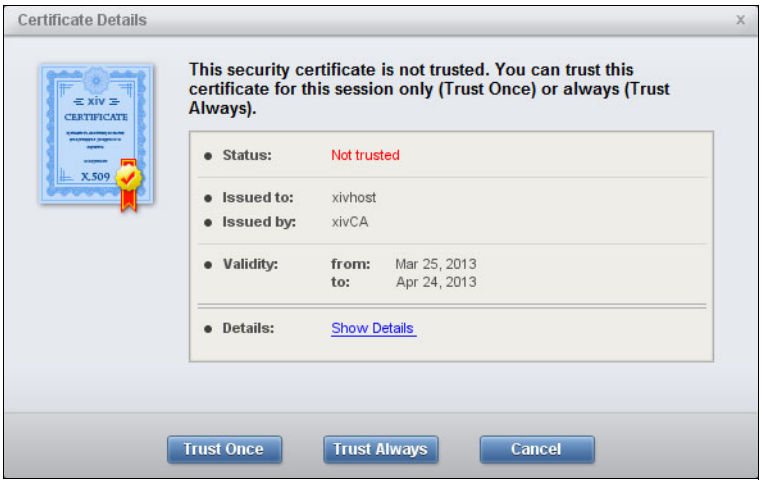

Certificate Details
This security certificate is not trusted. You can trust this certificate for this session only (Trust Once) or always (Trust Always).
Status: Not trusted
Issued to: xivhost
Issued by: xivCA
Validity: from: Mar 25, 2013
to: Apr 24, 2013
Details: Show Details
Trust Once
Trust Always
Cancel

Figure 5-10 Certificate Details window

The certificate error is resolved and you are now able to securely connect to this system from the GUI using your own x509 certificate.

Note: The XIVTop tool uses the same certificate store as the XIV GUI. Trusting a certificate in the GUI means that the XIVTop tool will also trust that certificate.

5.3 Configuring IPv6 addresses

The use of IPv6 addressing for XIV Storage System management and virtual private network (VPN) ports is supported by XIV Storage System Software Version 11.1 and later. To use IPv6 addressing, you must first enable it on the XIV Storage System.

5.3.1 Enabling IPv6 support

In the GUI, navigate to the System view of the chosen XIV Storage System and click **Systems** → **System Settings** → **System** to open the System panel. Then, click the **Parameters** tab, as shown in Figure 5-11.

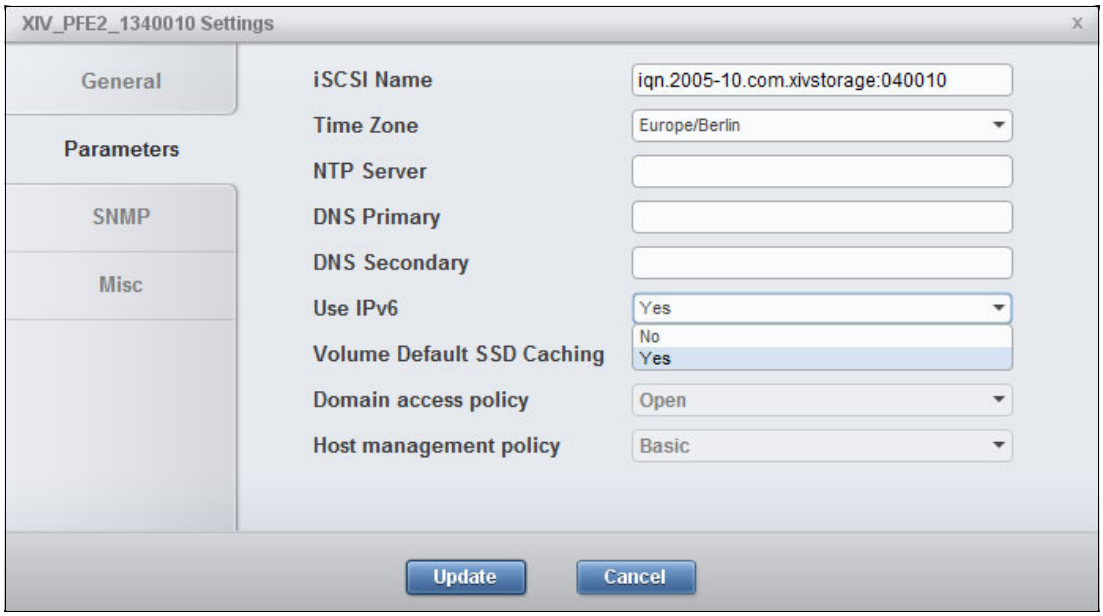


Figure 5-11 Enable IPv6

In the drop-down box for Use IPv6, select **Yes** and click **Update**.

5.3.2 Configuring IPv6 addresses

In the main window, click the **Backview** arrow next to the XIV System Storage image, as shown in Figure 5-12 on page 216, to open the view of the system ports.

Figure 5-12 GUI Backview arrow

Configuring management port IPv6 addresses

To configure IPv6 addresses on a management port, right-click that port and select **Update Management IP Interface**, as shown in Figure 5-13.

Figure 5-13 Update Management IP Interface

Enter the corresponding three IPv6 addresses in the fields shown in Figure 5-14, and click **Update**.

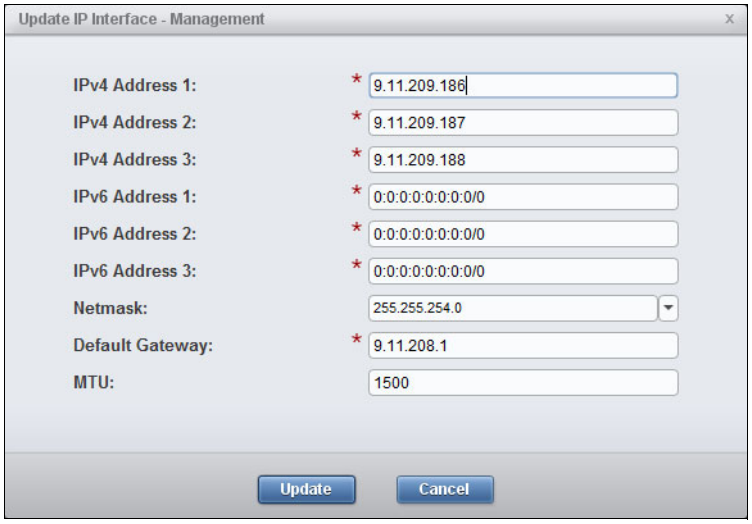


Figure 5-14 Update management port IPv6 addresses

Configuring VPN port IPv6 addresses

To configure IPv6 addresses on a VPN port, right-click that port and select **Update VPN IP Interface**, as shown in Figure 5-15.

Figure 5-15 Update VPN Interface

Enter the corresponding two IPv6 addresses in the fields shown in Figure 5-16, and click **Update**.

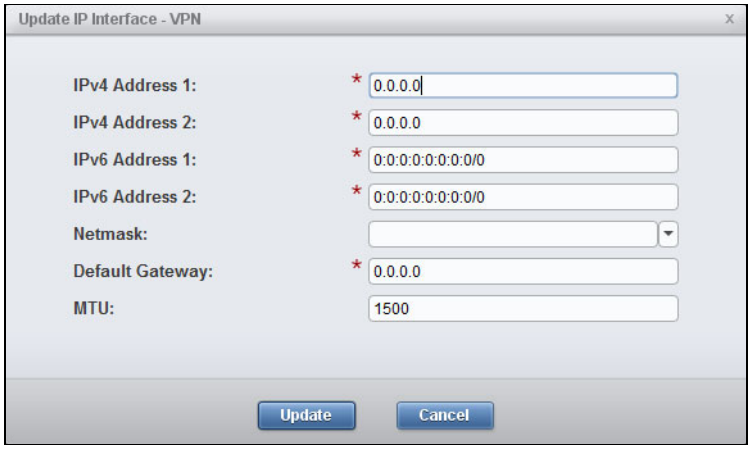


Figure 5-16 Update VPN port IPv6 addresses

5.4 Configuring Internet Protocol Security connectivity

Internet Protocol Security (IPSec) is a protocol suite that allows for enhanced security of IP communications through the authentication and encryption of IP packets. The XIV Storage System Software and XIV Management Tools allow for the use of passkey or certificate authentication to establish IPSec connectivity between management workstations and the management or VPN ports of XIV systems.

To configure IPSec connectivity between a management workstation and an XIV Storage System, perform the following steps:

1. In the GUI, navigate to the System view of the chosen XIV Storage System and click **Systems** → **System Settings** → **IPSec** to open the IPSec panel, as shown in Figure 5-17 on page 218.

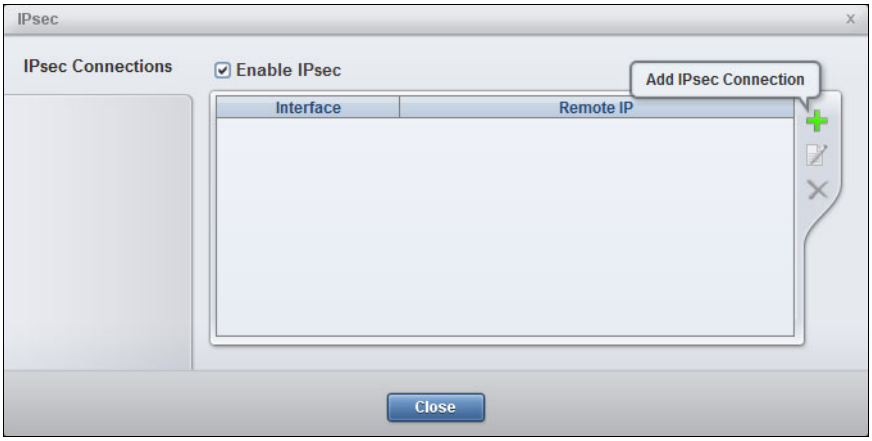


Figure 5-17 IPSec configuration

2. Click the green plus (**+**) icon on the right to open the Add IPSec Connection panel, which is shown in Figure 5-18.

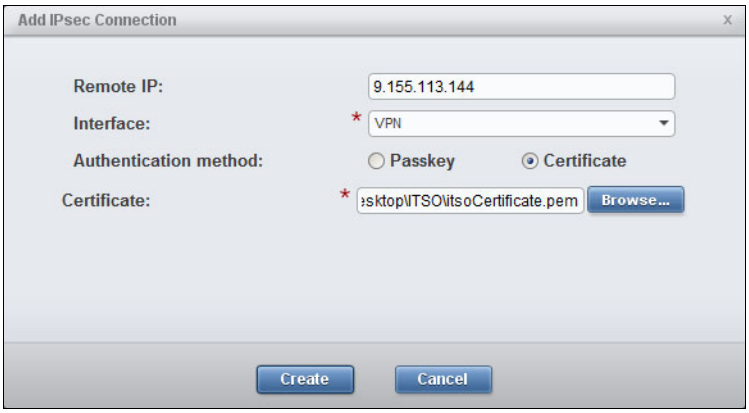


Figure 5-18 Add IPSec Connection with certificate authentication

In the Remote IP field, enter the IP address of the remote workstation you will use to connect over IPSec.

Choose which XIV Storage System interface type on which you want to establish the IPSec connections. The available choices are Management and VPN.

Select the radio button for the authentication method that you want to use:

– For Certificate authentication, click **Browse** to select and upload your certificate file. For more information about obtaining certificates, see 5.2.1, “Managing x509 certificates with the XIV Storage Management GUI” on page 209, and “Certificate authority setup” on page 460.

– For Passkey authentication, enter the passkey to be used by the host to authenticate in the Passkey field, as shown in Figure 5-19. Repeat the passkey in the Passkey Verify field.

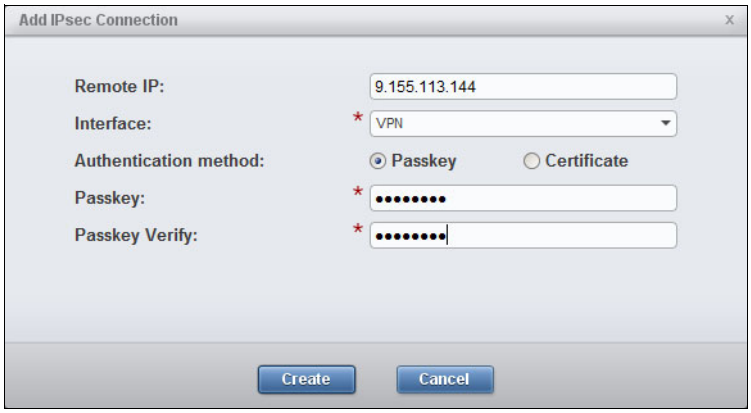


Figure 5-19 Add IPSec Connection with passkey authentication

3. Click **Create** to create the IPSec connection on the XIV System. The completed connection is shown in Figure 5-20.

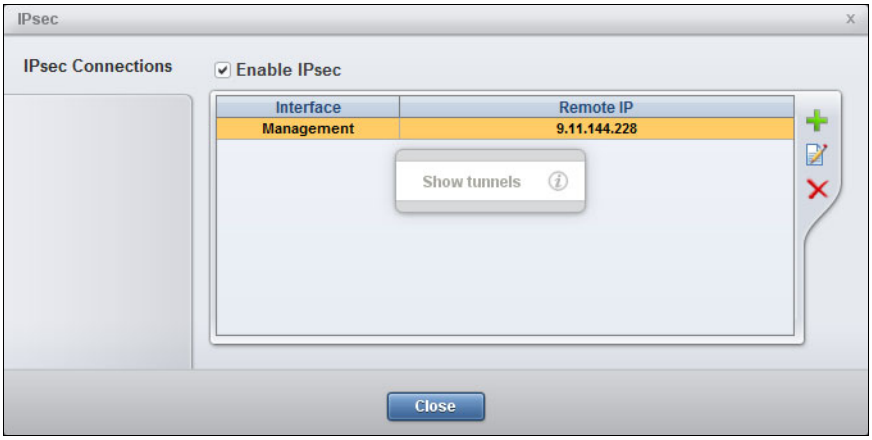


Figure 5-20 Completed IPSec connection configuration

4. When an IPSec connection is established from the remote IP address, connection details can be viewed by right-clicking the connection entry and choosing **Show tunnels**, as shown in Figure 5-20.

5.5 Native user authentication

To prevent unauthorized access to the configuration of the storage system and ultimately to the information stored on its volumes, the XIV Storage System uses password-based user authentication. Password-based authentication is a form of challenge-response authentication protocol where the authenticity of a user is established by presenting that user with a question “challenge” and comparing the answer “response” with information stored in a credential repository.

Note: If you plan to use the Hyper-Scale Manager, see the IBM Redpaper, *IBM Hyper-Scale for the XIV Storage System*, REDP-5053, and 5.5.3, “Security considerations when using Hyper-Scale Manager” on page 232.

5.5.1 XIV Storage System local credential repository

By default, the XIV Storage System is configured to use native (XIV managed) user authentication. Native user authentication uses the credential repository stored locally on the XIV Storage System. The XIV local credential repository maintains the following information:

- User name
- User password
- User role
- User group
- Optional account attributes

User name

A user name is a string of 1 - 63 characters that can contain only `a - z`, `A - Z`, `0 - 9`, `.-_~`, and `space` symbols. User names are case-sensitive. The XIV Storage System is configured with a set of predefined user accounts. Predefined user names and corresponding default passwords exist to provide initial access to the XIV Storage System at the time of installation, for system maintenance, and for integration with applications such as the IBM Tivoli Storage Productivity Center.

The following user accounts are predefined on the XIV Storage System:

- Technician

  This account is used by the IBM Service Support Representative (SSR) to install the XIV Storage System.

- Admin

  This account provides the highest level of client access to the system. It can be used for creating new users and changing passwords for existing users in native authentication mode.

  Important: Use of the *admin* account should be limited to the initial configuration when no other user accounts are available. Access to the admin account needs to be restricted and securely protected.

- xiv_development and xiv_maintenance user

  These IDs are special case predefined internal IDs that can only be accessed by qualified IBM development personnel and SSRs.

Predefined user accounts cannot be deleted from the system and are always authenticated natively by the XIV Storage System even if the system operates under Lightweight Directory Access Protocol (LDAP) authentication mode.

User accounts can initially be created by the admin user only. After the admin user creates a user account and assigns it to the storageadmin (Storage Administrator) role, then other user accounts can be created by this storageadmin user.

In native authentication mode, the system is limited to creating up to 128 user accounts. This number includes the predefined users.

User password

The *user password* is a secret word or phrase used by the account owner to gain access to the system. The user password is used at the time of authentication to establish the identity of that user. User passwords can be 6 - 12 characters long and may include these characters:

a - z, A - Z, ~, !, @, #, $, %, ^, &, *, (,), _ , + , -, =, {, }, |, :, ;, <, >, ?, ., /, \, [, and]

The passwords must not have spaces between characters. In native authentication mode, the XIV Storage System verifies the validity of a password when the password is assigned.

Predefined users have the default passwords shown in Table 5-1 assigned at the time of XIV Storage System installation.

Table 5-1 Default passwords

Predefined user	Default password
admin	adminadmin
technician	Predefined. Used only by the XIV Storage System technicians.
xiv_development	Predefined. Used only by the IBM XIV development team.
xiv_maintenance	Predefined. Used only by the IBM XIV maintenance team.

Important: The default `admin` password must be changed at the time of installation to prevent unauthorized access to the system. For instructions, see “Adding users with the XIV Storage Management GUI” on page 225.

The following restrictions apply when working with passwords in native authentication mode:

- For security purposes, passwords are not shown in user lists.
- Passwords are user changeable. Users can change only their own passwords.
- Only the predefined user *admin* can change the passwords of other users.
- Passwords are changeable from both the XCLI and the XIV Storage Management GUI.
- Passwords are case-sensitive.
- User password assignment is mandatory at the time a new user account is created.
- Creating user accounts with an empty password or removing a password from an existing user account is not permitted.

User roles

There are nine predefined user roles (in the XIV Storage Management GUI and the XCLI). Roles are referred to as *categories* and are used for day-to-day operation of the XIV Storage System. The first four categories listed (storageadmin, applicationadmin, securityadmin, and read only) are allowed to have multiple users created with these roles; the other four are preassigned by the system and do not allow additional users to be created with these roles. The following section describes predefined roles, their level of access, and applicable use:

- storageadmin

  The storageadmin (Storage Administrator) role is the user role with highest level of access available on the system. A user assigned to this role can perform changes on any system resource except for maintenance of physical components, changing the status of physical components, or effecting changes related to encryption.

- applicationadmin

  The *applicationadmin* (application administrator) role provides flexible access control over volume snapshots. A user assigned to the applicationadmin role can create snapshots of volumes, perform mapping of their own snapshots to assigned hosts, and delete their own snapshots. The user group to which an application administrator belongs determines the hosts and their respective set of mapped volumes that the application administrator is allowed to manage. If a user group is defined with `access_all="yes"`, application administrators who are members of that group can manage all volumes on the system. For more details about user group membership and group-to-host association, see "User groups" on page 224.

- securityadmin

  The securityadmin (security administrator) role is to manage certificates, key servers, and other items related to encryption. Defining at least two security administrators is a must, because that is the minimum number required to enable recovery keys. For details about encryption with XIV, see the IBM Redpaper titled *XIBM XIV Security with Data-at-Rest Encryption*, REDP-5047.

- readonly

  As the name implies, users assigned to the readonly role can only view system information. A typical use for the readonly role is a user who is responsible for monitoring system status, system reporting, and message logging. This user must not be permitted to make changes on the system.

- technician

  The technician role has a single predefined user name (technician) assigned to it, and it is intended to be used by IBM support personnel for maintaining the physical components of the system. The technician is limited to the following tasks: physical system maintenance and phasing components in or out of service. The technician has restricted access to the system and is unable to perform any configuration changes to pools, volumes, or host definitions on the XIV Storage System.

- xiv_development

  The xiv_development role has this single predefined user name assigned to it and it is intended to be used by IBM development personnel.

- xiv_maintenance

  The xiv_maintenance role has this single predefined user name assigned to it and it is intended to be used by IBM maintenance personnel.

- xiv_hostprofiler

  The xiv_hostprofiler role has this single predefined user name assigned to it and it is intended to be used for gathering additional information about hosts attached to the XIV.

User roles: There is no capability to add new user roles or to modify predefined roles. In native authentication mode, after a user is assigned a role, the only way to assign a new role is to first delete the user account and then re-create it.

Table 5-2 shows the predefined users and their roles.

Table 5-2 Predefined user role assignment

Predefined user	User role
admin	storageadmin
technician	technician
xiv_development	xiv_development
xiv_maintenance	xiv_maintenance
xiv_hostprofiler	xiv_hostprofiler
xiv_msms	storageadmin

Native authentication mode implements user role mechanism as a form of *role-based access control* (RBAC). Each predefined user role determines the level of system access and associated functions that a user is allowed to use.

RBAC: The XIV Storage System implements role-based access control (RBAC) based authentication and authorization mechanisms.

All user accounts must be assigned to a single user role. Assignment to multiple roles is not permitted. Deleting or modifying role assignment of natively authenticated users is also not permitted.

Global administrator and domain administrator

The XIV Storage Software Version 11.5 introduced the domain-based multitenancy feature, which is described in 2.4.3, “Multitenancy” on page 30.

When you add a user, the user can be assigned to zero (no domain), one, or all domains defined on that XIV system, as shown in Figure 5-21.

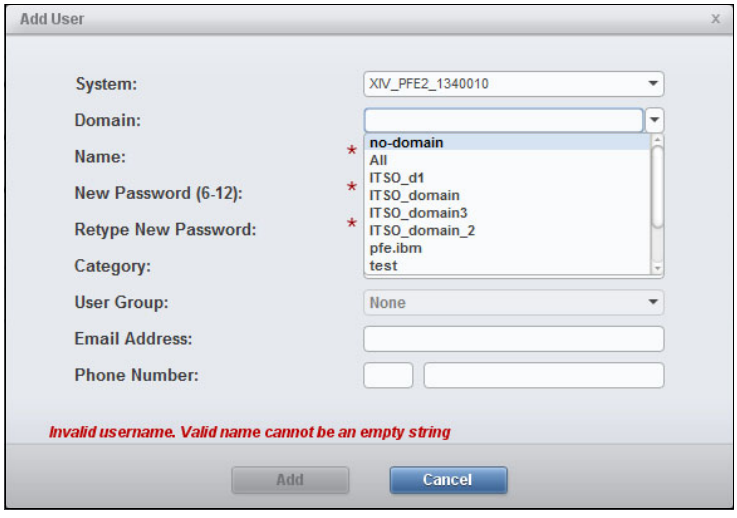


Figure 5-21 Add User to domain

If a particular user needs to be assigned to multiple domains, the additional association must be made from the Domains view, as explained under “Managing domain user associations” on page 189.

With the introduction of domains, there are two levels of authority for a user in a specific role:

- Domain administrator

  A user who is associated with one or more domains is a domain administrator and has access rights (as implied by the user's role, such as readonly or storageadmin) to the entities uniquely associated with the pertinent domains.

- Global administrator

  A user who is not associated with any domain is a global administrator and has access rights to all of the entities that are not uniquely associated with a domain.

  A global administrator who is a storageadmin can create, edit, and delete domains and can associate resources with a domain.

  An open or closed policy can be defined so that a global administrator may, or may not, be able to extend the reach into a domain

  Intervention of a global administrator, that has permissions for the global resources of the system, is only needed for these situations:

  – Initial creation of the domain and assigning a domain administrator
  – Resolving hardware issues

For more information about managing domains, see 4.8, "Multitenancy" on page 174.

User groups

A user group is a group of application administrators who share a set of snapshot creation permissions. The permissions are enforced by associating the user groups with hosts or clusters. User groups have the following characteristics:

- Only users assigned to the applicationadmin role can be members of a user group.
- A user can be a member of a single user group.
- A maximum of eight user groups can be created.
- Group names are case-sensitive.
- In native authentication mode, a user group can contain up to eight members.
- If a user group is defined with `access_all=”yes”`, users assigned to the applicationadmin role who are members of that group can manage all snapshots on the system.
- A user must be assigned to the storageadmin role to be permitted to create and manage user groups.

Important: A user group membership can be defined only for users assigned to the applicationadmin role.

User group and host associations

Hosts and clusters can be associated with only a single user group. When a user is a member of a user group that is associated with a host, that user is allowed to manage snapshots of the volumes mapped to that host.

User group and host associations have the following properties:

- User groups can be associated with both hosts and clusters. This configuration enables limiting group member access to specific volumes.
- A host that is part of a cluster can be associated only with a user group through user group to cluster association. Any attempts to create user group association for that host fail.

- When a host is added to a cluster, the association of that host is removed. Limitations on the management of volumes mapped to the host is controlled by the association of the cluster.
- When a host is removed from a cluster, the association of that cluster remains unchanged. This configuration enables continuity of operations so that all scripts relying on this association continue to work.

Optional account attributes

In this section, we describe optional attributes for email and phone numbers:

- Email: Email is used to notify specific users about events through email messages. Email addresses must follow standard formatting procedures.

  Acceptable value: Any valid email address. A default value is not defined.

- Phone and area code: Phone numbers are used to send SMS messages to notify specific users about system events. Phone numbers and area codes can be a maximum of 63 digits, hyphens (-), and periods (.)

  Acceptable value: Any valid telephone number. A default value is not defined.

5.5.2 Managing user accounts with the XIV Storage Management GUI

This section illustrates the use of the XIV Storage Management GUI in native authentication mode for creating and managing user accounts, as well as for creating user groups and defining group membership.

Adding users with the XIV Storage Management GUI

The following steps require that you initially log on to the XIV Storage System with storage administrator access rights (storageadmin role). If this is the first time that you are accessing the system, use the predefined user admin (default password adminadmin).

Follow these steps:

1. Open the XIV Storage Management GUI and log on as shown in Figure 5-22.

XIV Storage Management
User:
admin
Password:
Mode:
Direct
Login
Licensed Materials - Property of IBM Corporation and other(s). Copyright © 2008, 2012. IBM and XIV are a registered trademarks of IBM Corporation in the United States, other countries, or both.
IBM

Figure 5-22 XIV Storage Management GUI login

2. Users are defined per system. If you manage multiple systems and they have been added to the XIV Storage Management GUI, select the system that you want to work with.

3. In the main Storage Manager GUI window, hover the cursor over the padlock icon to display the Access menu. All user access operations can be performed from the Access menu (see Figure 5-23). In the Version 3 XIV Storage Management GUI, only the Users menu option shows. Click **Users**, as shown in Figure 5-23

Figure 5-23 XIV Storage Management GUI Access menu

4. By default, the Ungrouped group is displayed, and is shown in a collapsed state. Click the plus symbol (**+**) to expand its state.

   If the storage system is being accessed for the first time, the window shows the predefined users only, as shown in Figure 5-24. The default columns are Name, Category, Phone, Email, and Group.

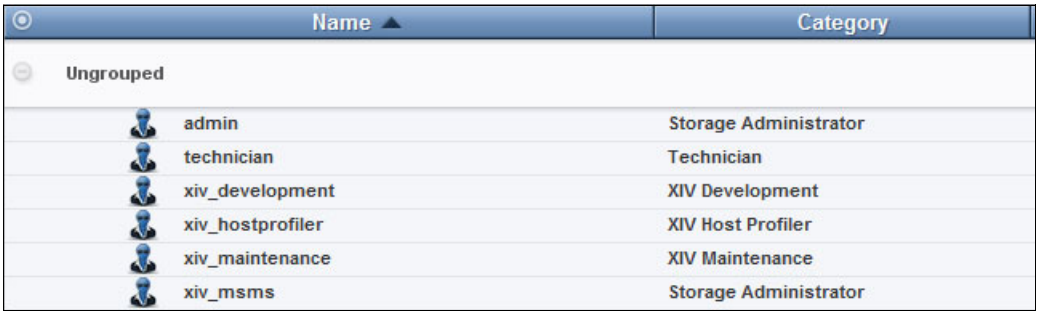


Figure 5-24 XIV Storage Management GUI users management

5. Change the default password for the admin user, which can be accomplished by right-clicking the user name and selecting **Change Password** from the menu, as shown in Figure 5-25. You must be logged in as admin to change the admin password.

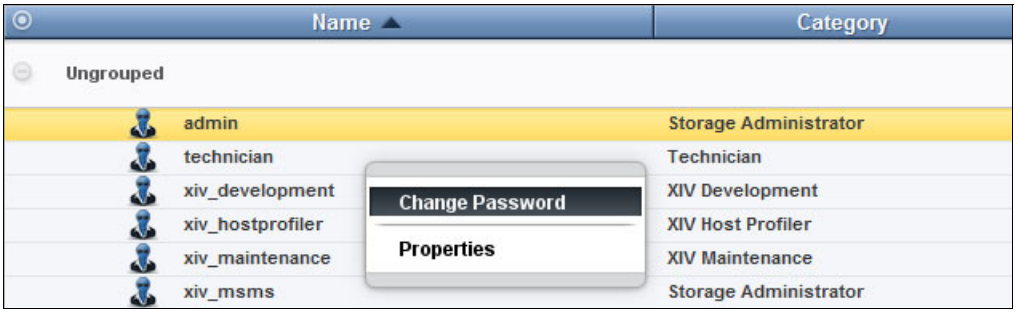


Figure 5-25 XIV Storage Management GUI admin user change password

6. To add a user, you can either click the **Add User** icon in the menu bar or right-click the empty space to get the menu, as shown in Figure 5-26. Click **Add User**. A maximum of 128 users are possible in the regular XIV access control.

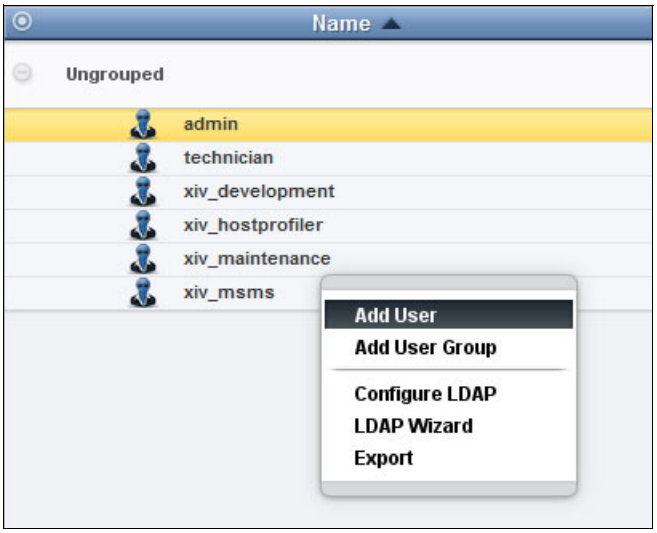


Figure 5-26 XIV Storage Management GUI Add User option

7. The Add User window opens, as shown in Figure 5-27. A user is defined by a unique name and a password. The default role (denoted as Category in the window) is *Storage Administrator*. A category must be assigned. Optionally, enter the email address and phone number for the user. Click **Add** to create the user and return to the Users window.

Note: Additional field called Domain is presented only if domain is already defined to the XIV system (see Figure 5-27) otherwise this field is hidden. A user not associated with any domain is a global user. See 4.8.2, “Managing domains with the XIV Storage Management GUI” on page 175.

Figure 5-27 XIV Storage Management GUI Add User attributes

8. If you need to test the user that you defined, click the current user name shown in the upper right of the IBM XIV Storage Management window (Figure 5-28), and log in with the new user's credentials.

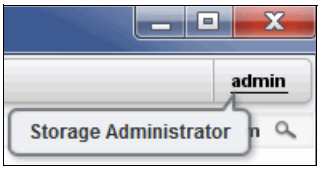


Figure 5-28 XIV Storage Management GUI quick user change

Defining user groups with the XIV Storage Management GUI

The XIV Storage System can simplify various user management tasks if you use the capability to create user groups.

Groups: User groups apply only to users that are assigned to the applicationadmin role.

A user group can also be associated with one or multiple hosts or clusters.

The following steps illustrate how to create user groups, add users (with application administrator role) to the group, and how to define host associations for the group:

1. Ensure that you are logged in as admin (or another user with storage administrator rights) and in the Users window. In our scenario, we create a user group called `Application01_Group`.
2. To add a user group, either click the **Add User Group** icon (shown in Figure 5-29), or right-click in an empty area of the User Group table and select **Add User Group** from the menu, as shown in Figure 5-29.

Figure 5-29 Add User Group

3. From the Add User Group window that opens, enter a meaningful group name and then click **Add**.

Note: Additional field called Domain is presented only if domain is already defined to the XIV system (see Figure 5-30) otherwise this field is hidden.

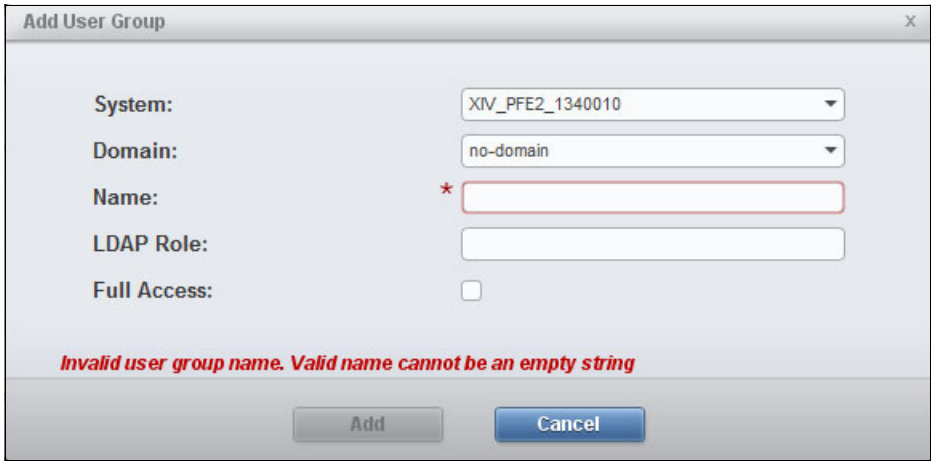


Figure 5-30 Enter new user group name

LDAP role: The LDAP Role field is not applicable to user group definition in native authentication mode and has no effect, even if a value is entered.

If a user group has the Full Access flag turned on, all members of that group have unrestricted access to all snapshots on the system.

At this stage, the user group `Application01_Group` still has no members and no associations defined. Next, we create an association between a host and the user group.

4. Right-click the name of the user group that you created to open a menu, and select **Update Access Control**, as shown in Figure 5-31.

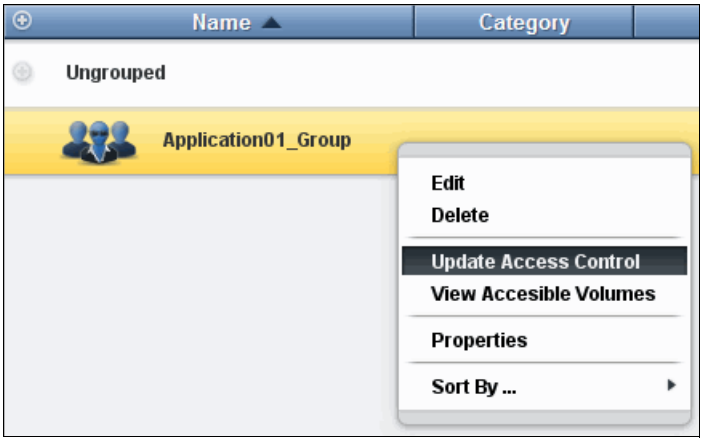


Figure 5-31 Update Access Control for a user group

The User Group Access Control window shown in Figure 5-32 on page 230 opens. The window contains the names of all the hosts and clusters defined to the XIV Storage System. The left pane shows the list of unauthorized hosts and clusters for this user group, and the right pane shows the list of hosts that have already been associated with the user group.

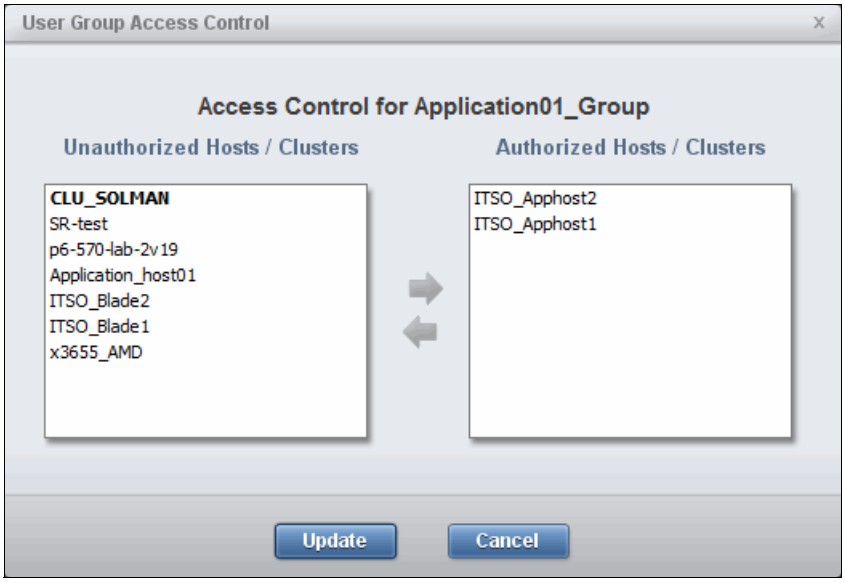


Figure 5-32 Access Control Definitions window

5. Add or remove hosts from either list by selecting a host and clicking the appropriate arrow. Finally, click **Update** to save the changes.

   You can verify which group that hosts have been associated with by viewing it in the Hosts and Clusters window. See Figure 5-33.

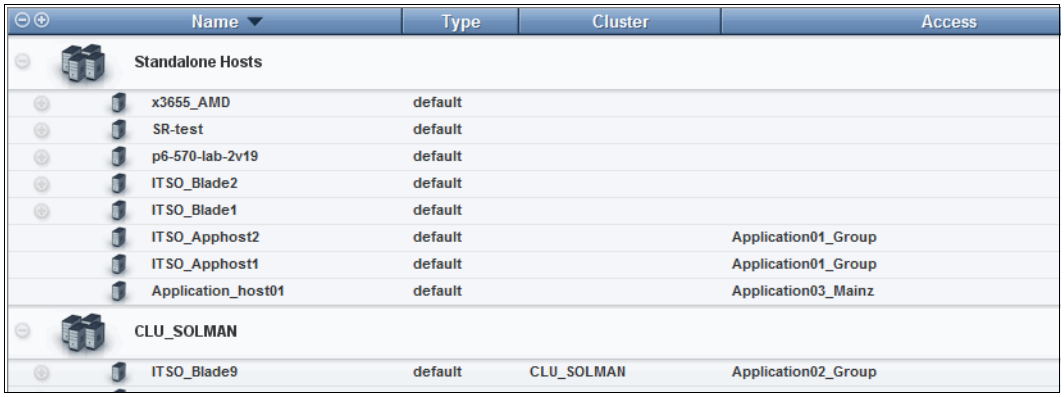


Figure 5-33 Hosts and groups associations

After a host (or multiple hosts) has been associated with a user group, you can define user membership for the user group (a user must have the application administrator role to be added to a user group).

6. Go to the Users window and right-click the user name to display the menu. From the menu (see Figure 5-34), select **Add to Group** to add this user to a group.

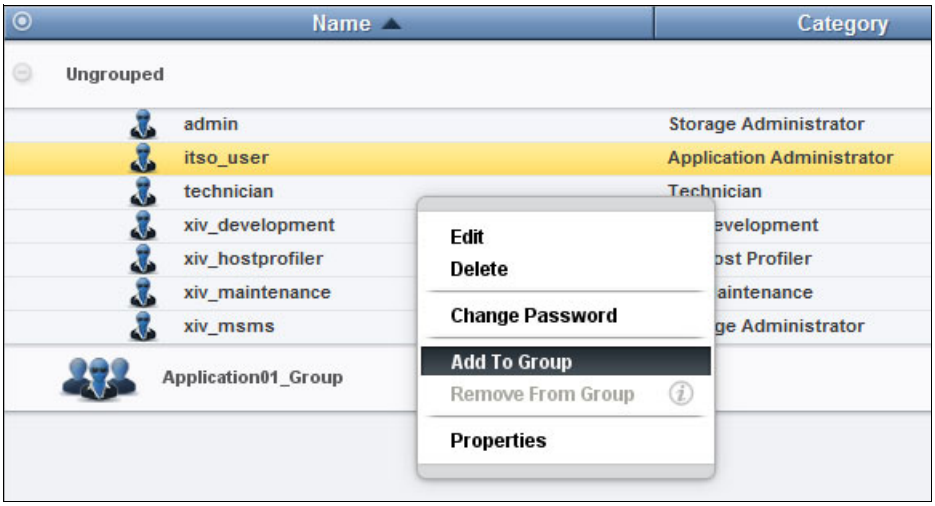


Figure 5-34 Add a user to a group

7. From the Select User Group window that opens, select the group that you want from the drop-down list and click **OK** (see Figure 5-35).

Figure 5-35 Select User Group window

The lab_admin user has been added as a member to the Application01_Group user group in this example.

8. Verify this group membership under the Application01_Group section of the Users window, as shown in Figure 5-36.

Figure 5-36 View user group membership

The lab_admin user is an applicationadmin with the Full Access right set to `no`. This user can now perform snapshots of volumes mapped to the hosts `ITSO_Apphost1` and `ITSO_Apphost2`. The user can also map, unmap, and delete snapshots from the application group.

5.5.3 Security considerations when using Hyper-Scale Manager

Hyper-Scale Manager allows access and distributes tasks among the following users:

- Root user

  The root user is an OS super user that carries out administrative tasks.

- Server admin user

  The server admin carries out inventory configuration tasks.

- Maintenance user

  The maintenance user carries out Hyper-Scale Manager configuration tasks. This user is the client's way to interact with any of the Hyper-Scale Manager operations by using Secure File Transfer Protocol (SFTP), such as setting logs, copying the upgrade file, and getting backups.

Note: This section is meant as information only. For more assistance and directions, see the *IBM XIV Storage System Management Tools Version 4.0 User Guide*, SC27-4230-00.

Root user operations

The menu that is shown in Figure 5-37 on page 233 list various configuration options for the Hyper-Scale Manager tool, which is available for the root user. Security options are under 5 and 6. The Hyper-Scale Manager user must log in to the Hyper-Scale Manager with *root credentials* to access the menu items. See *IBM XIV Storage System Management Tools Version 4.0 User Guide*, SC27-4230-00, for instructions.

The LDAP server is set up during the installation of Hyper-Scale Manager. See 4.3, "XIV Storage Management software use" on page 117.

```
------------------------------------------------------------------
------------ IBM Hyper-Scale Manager v1.5.0.21 ---------------
------------------------------------------------------------------

------------------------26/06/2014 15:22-------------------------

Super User Menu
---------------
IBM Hyper-Scale Manager is running

 1) Server Configuration Options
 2) Gather Manager Logs
 3) Show Audit Log
 4) Show Upgrade Logs
 5) Manager Inventory Options
 6) Backup/Restore Options
 7) Change 'root' Password
 8) Change 'maintenance' Password
 9) Change Manager Access Code
10) Start/Stop Hyper-Scale Manager Service
11) Replace Hyper-Scale Manager Certificate
12) Change XIV GUI Communication Ports
13) Change RESTful Interface Port
14) Upgrade Hyper-Scale Manager
15) Reboot
16) Shutdown
17) Exit
Your Selection>
```

Figure 5-37 Root user menu window

Tip: A good practice is to configure at least one admin-level credential (user ID and password) to be the same as the MSMS login credentials (user ID and password).

Server admin operations

The menu that is shown in Figure 5-38 on page 234 is the Manager Inventory Options menu. Using this menu, the administrator can manage the inventory of XIV systems monitored by the Hyper-Scale Manager and control their status.

From a security standpoint, the following options are available:

- Reauthenticate all users.

  This option forces the reauthentication of all GUI logged in users on all XIV. This option can be used when there is an authentication change for users and you want it to take effect immediately.

- Changing the System Machine Account password.

  The System Machine Account is a system user that monitors the system on the Hyper-Scale Manager. The password can be changed either from the root menu or via the GUI. In addition, the password must be changed on all systems.

- Changing the server admin password.

  This menu option changes the System Machine Account password.

  See the *IBM XIV Storage System Management Tools Version 4.0 User Guide*, SC27-4230-00, for instructions.

```
------------------------------------------------------------------
------------ IBM Hyper-Scale Manager v1.5.0.21 ---------------
------------------------------------------------------------------

------------------------26/06/2013 15:46-------------------------

Manager Inventory Options
-------------------------
IBM Hyper-Scale Manager is running.

 1) System Inventory List
 2) Add System
 3) Modify System
 4) Remove System
 5) System Monitoring Suspend
 6) System Monitoring Resume
 7) Re-authenticate All Users
 8) Change System Machine Account
 9) Manage System Certificates
10) Manage Capacity Planning Data
11) Back to previous menu
Your Selection>
```

Figure 5-38 Manager Inventory Options menu

Maintenance user operations

Accessing the MSMS with *maintenance credentials* allows you to perform various maintenance tasks, such as gathering server logs, backing up or restoring the database, replacing the security certificate, or upgrading the MSM.

See *IBM XIV Storage System Management Tools Version 4.0 User Guide* SC27-4230-00 for more information.

5.5.4 Managing user accounts using the XCLI

This section summarizes the commands and options available to manage user accounts, user roles, user groups, group memberships, and user group to host associations through the XCLI Session.

Table 5-3 on page 235 shows the various commands and a brief description for each command. The table also indicates the user role that is required to issue specific commands. New access control commands come with XIV code release 11.5 for domain-based multitenancy users.

Important: With XIV Storage Software v11.4, a new *secadmin* role is introduced and required to execute specific commands related to encryption management. Details are covered in the IBM Redpaper titled *XIV Security with Data-at-Rest Encryption*, REDP-5047.

Table 5-3 XCLI access control commands

Command	Description	Role required to use command
access_define	Defines an association between a user group and a host.	storageadmin
access_delete	Deletes an access control definition.	storageadmin
access_list	Lists access control definitions.	storageadmin, readonly, and applicationadmin
user_define	Defines a new user.	storageadmin and securityadmin
user_delete	Deletes a user.	storageadmin and securityadmin
user_list	Lists all users or a specific user.	storageadmin, readonly, securityadmin, applicationadmin
user_rename	Renames a user.	storageadmin and securityadmin
user_update	Updates a user. You can rename the user, change a password, modify the Access All setting, modify email, area code, or phone number.	storageadmin, applicationadmin, securityadmin
user_group_add_user	Adds a user to a user group.	storageadmin
user_group_create	Creates a user group.	storageadmin
user_group_delete	Deletes a user group.	storageadmin
user_group_list	Lists all user groups or a specific one.	storageadmin, readonly, and applicationadmin
user_group_remove_user	Removes a user from a user group.	storageadmin
user_group_rename	Renames a user group.	storageadmin
user_group_update	Updates a user group.	storageadmin
domain_list_users	Lists users associated with domains.	storageadmin, securityadmin, applicationadmin, readonly, technician
domain_add_user	Associate a user with a domain.	storageadmin, securityadmin
domain_remove_user	Disassociate a user from a domain.	storageadmin,securityadmin

Adding users with the XCLI

Before performing the following steps, the XCLI component must be installed on the management workstation, and a storageadmin user is required. For more information about installing the XIV Management software, including the XCLI, see 4.2, “XIV Storage Management software installation” on page 114.The following examples assume a Windows management workstation.

To add users, complete the following steps:

1. Use the **user_list** command to obtain the list of predefined users and categories, as shown in Example 5-1. This example assumes that no users, other than the default users, have been added to the system.

Example 5-1 XCLI user_list

```
>>user_list
Name              Category          Group                Active...  Access All
xiv_development   xiv_development                        yes
xiv_maintenance   xiv_maintenance                        yes
admin             storageadmin                           yes
technician        technician                             yes
espadmin          storageadmin                           yes
itso              storageadmin                           yes
lab_admin         applicationadmin  Application01_Group  yes   ...  no
```

2. If this system is a new system, change the default password for the admin user by running **update_user**, as shown in Example 5-2.

Example 5-2 XCLI user_update

```
>>user_update user=admin password=New-Passw0rd password_verify=New-Passw0rd
Command executed successfully.
```

3. Add a user, as shown in Example 5-3. Define a user using a unique name, password, and role (designated here as *category*).

Example 5-3 XCLI user_define

```
>>user_define user=lab_user password=XIV-R3 password_verify=XIV-R3
category=applicationadmin
Command executed successfully.
```

To add user to domain, complete the following steps:

1. Use the **domain_list_users** command to obtain the list of predefined and domain associated users and categories, as shown in Example 5-4. This example is based on the assumption that a domain is already defined and one admin associated user is already present in the system.

Example 5-4 XCLI domain_list_user

```
XIV_PFE2_1340010>>domain_list_users
Domain      User                    Category
ITSO_d1     ITSO_User1              storageadmin
no-domain   Residency               applicationadmin
no-domain   xiv_development         xiv_development
no-domain   xiv_maintenance         xiv_maintenance
no-domain   admin                   storageadmin
no-domain   technician              technician
no-domain   xiv_hostprofiler        xiv_hostprofiler
```

2. Add the same user to another domain, as shown in Example 5-5 on page 237.The example assumes that another domain is already defined. Specify the domain and user name. A user can have a domain administrator role in more than one domain.

Example 5-5 XCLI domain_add_user

```
XIV_PFE2_1340010>>domain_add_user domain=pfe.ibm user=ITSO_User1
Command executed successfully.
```

Defining user groups with the XCLI

To define a user group, complete the following steps:

1. Run **user_group_create**, as shown in Example 5-6, to create a user group called Application03_Mainz.

   Example 5-6 XCLI user_group_create

   ```
   >>user_group_create user_group=Application03_Mainz
   Command executed successfully.
   ```

   Spaces: Avoid spaces in user group names. If spaces are required, the group name must be placed between quotation marks, such as “name with spaces.”

   The Application03_Mainz user group is empty and has no associated host. The next step is to associate a host or cluster.

2. Associate the application_Mainz user group to the Application_host01 host, as shown in Example 5-7.

   Example 5-7 XCLI access_define

   ```
   >>access_define user_group=Application03_Mainz host=Application_host01
   Command executed successfully.
   ```

   A host has been assigned to the user group. The user group does not have any users included.

3. Add the first user, as shown in Example 5-8.

   Example 5-8 XCLI user_group_add_user

   ```
   >>user_group_add_user user_group=Application03_Mainz user=lab_user
   Command executed successfully.
   ```

The user lab_user has been assigned to the Application03_Mainz user group. This user is an applicationadmin with the Full Access right set to no.

5.5.5 Password management and resets

Password management in native authentication mode is internal to the XIV Storage System. The XIV Storage System has no built-in password management rules such as password expiration, preventing reuse of the same passwords, or password strength verification. Furthermore, if you want to log on to multiple systems at the same time through the XIV Storage Management GUI, your user ID must be registered with the same password on all XIV systems that are managed by the XIV Storage Management GUI.

In native authentication mode, if users can log in, they can change their own passwords.

The predefined admin user is the only user that is authorized to change other users' passwords. Direct access to a user credential repository is not permitted. System security is enforced by allowing password changes only through the XIV Storage Management GUI and XCLI.

Figure 5-39 shows how to change a password. Right-click the selected user in the Users window and click **Change Password** from the menu.

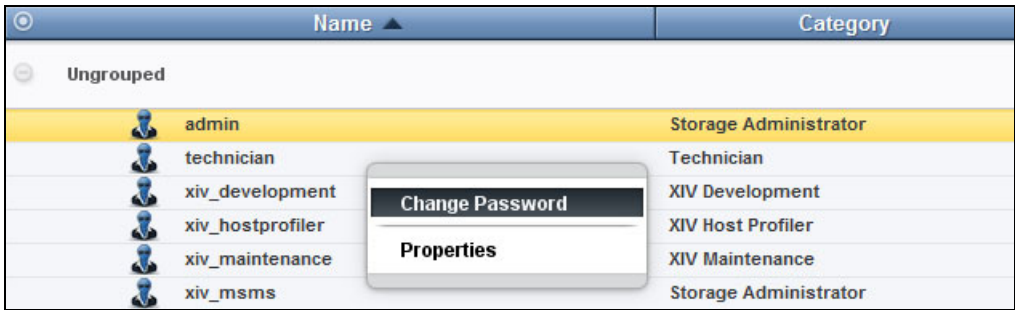


Figure 5-39 XIV Storage Management GUI Change Password menu

The Change Password window that is shown in Figure 5-40 opens. Enter the New Password and then retype it for verification in the appropriate field (only alphanumeric characters are allowed). Click **Update**.

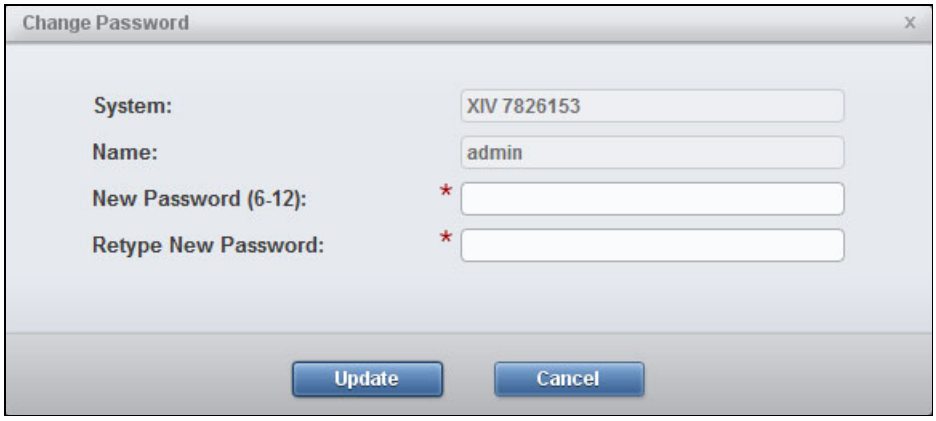


Figure 5-40 XIV GUI Change Password window

5.5.6 Managing multiple systems

Managing multiple XIV systems is straightforward in native authentication mode.

Authorization rule for managing multiple systems

Because user credentials are stored locally on every XIV Storage System, you must keep the same user name and password on separate XIV systems to allow for quick transitions between systems in the XIV Storage Management GUI. This approach is especially useful in Remote Mirror configurations, where the storage administrator is required to switch from source to target system.

Figure 5-41 shows the XIV Storage Management GUI view of multiple systems when using non-synchronized passwords. For this example, the systems named “XIV 1310133” and “XIV-02-131-114” have a user account named `lab_admin` that provides the storage admin level of access. Because the user ID is not configured for all XIV systems, only the two systems are currently shown as accessible.

The user can see the other system, but is unable to access it with the lab_admin user (the unauthorized system appears disabled). It also states that the user is unknown. If the system has the lab_admin defined with another password, the systems are shown in the same state.

Figure 5-41 shows a single user login view.

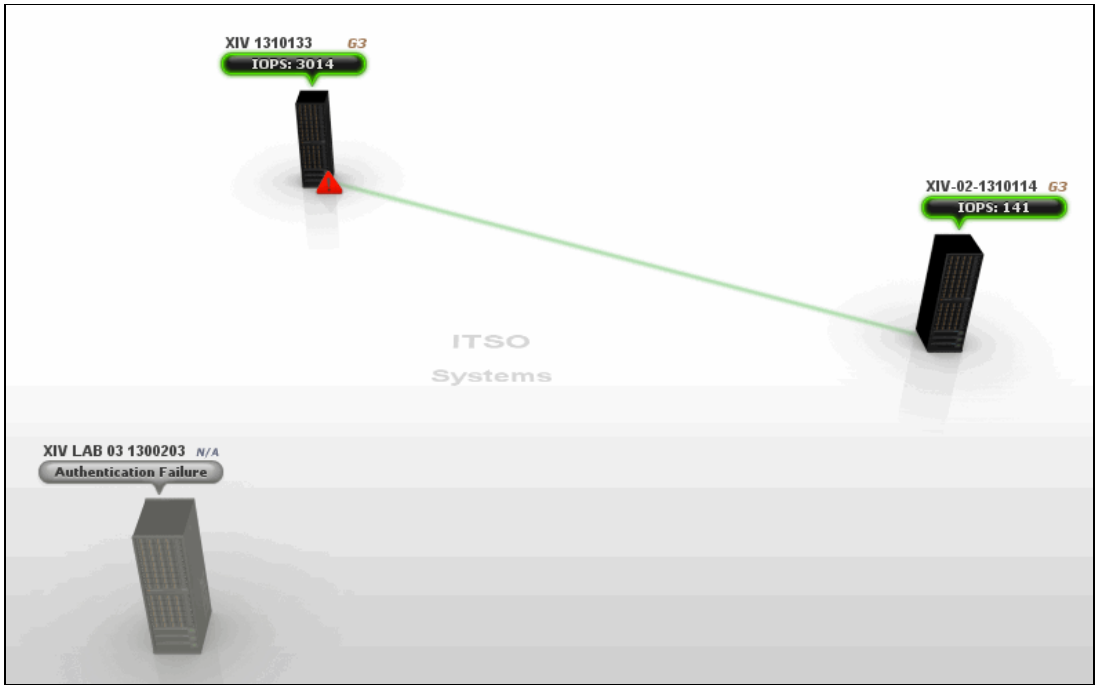


Figure 5-41 Single user login

To allow simultaneous access to multiple systems, the simplest approach is to have corresponding passwords manually synchronized among those systems. Figure 5-42 on page 240 illustrates the use of user account with passwords synchronized among four XIV systems. The storage administrator can easily switch between these systems for the activities without needing to log on each time with another password. Each XIV Storage System where the user was successfully authenticated is now displayed in color with an indication of its status.

Figure 5-42 Manual user password synchronization among multiple XIV systems

Direct mode GUI limitations for managing multiple systems

The XIV System Management GUI, Version 4.1 and higher, supports managing multiple XIV systems in a single GUI instance with the following limits:

- The maximum number of groups is 12.
- The maximum number of systems within a group is 12.
- The maximum number of managed systems is 144.

Important: You are advised to use Hyper-Scale Manager when managing more than 40 XIV systems. For more information, see the IBM Redpaper titled *IBM Hyper-Scale for the XIV Storage System*, REDP-5053.

It is possible to control more than 144 XIV Systems on a single workstation by creating another user profile, and starting another instance of the GUI:

- Using multiple profiles for accessing a different set of systems from the same workstation:
  - Access each profile by a command-line switch, where <user_dir> is the profile name:

    ```
    xivgui.exe -h <user_dir>
    ```

  - Multiple instances of the GUI can be opened, each with a different profile (Figure 5-43 on page 241), which can be useful for the following purposes:
    - Managing more than 144 XIV Systems
    - Managing different XIV systems using different credentials
  - Each profile has its own list of managed systems

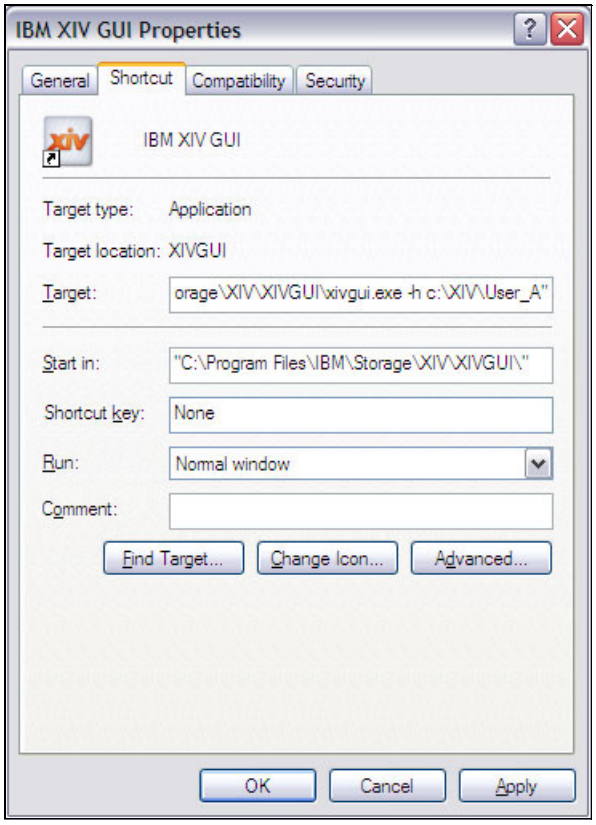


Figure 5-43 User profile for MultiSystem Manager

5.6 Enabling compliance with PCI-DSS

PCI -DSS (Payment Card Industry Data Security Standard) is the global data security standard adopted by the payment card brands for all entities that process, store, or transmit data.

The following security enhancements required by PCI-DSS standard are now enabled in XIV software Version 11.5, and enable clients to comply with PCI-DSS regulation and security requirements:

- Auditing via syslog
- Session Idle Timeout

These two features are in addition to other already available XIV features that all contribute to make XIV compliant with PCI-DSS, as summarized in Figure 5-44 on page 242.

PCI-DSS Section	Requirement	XIV Solution
2.3	Encrypt all non-console administrative access	Supported - SSL
3.1.1	Implement a data retention and disposal policy that includes ...Processes for secure deletion of data when no longer needed	Encryption disable/host level overwrite - out of XIV scope – disable encryption or change key – data is no longer available/readable
3.4.1, 3.5, 3.6	Disk encryption and key management requirements	SED feature of 11.4 supports these – key splitting – 2 key managers
8.5.9	Change user passwords at least every 90 days	Enforce password expiration – xiv outsources to LDAP server
8.5.10-.14	Minimum password length … passwords containing both numeric and alphabetic characters … Limit repeated access attempts … Set the lockout duration to a minimum of 30 minutes	LDAP server is a workaround for 8.5.9~8.5.14 except storage admin
8.5.15	If a session has been idle for more than 15 minutes, require the user to re-authenticate	Including GUI and XCLI
10.5.1-.7	Audit trails	Support via syslog server – outsource to syslog server

Figure 5-44 PCI-DSS support in XIV

5.6.1 Auditing via syslog

XIV gives you the ability to configure a syslog server for your XIV system. All user actions (including all administrative commands and parameters from the XCLI or GUI) performed on an XIV system are logged and exported to this server.

Syslog is commonly used for computer system management and security auditing, as well as generalized informational, analysis, and debugging messages. XIV now has the capability of auditing all user-entered commands.

Currently, the feature supports the definition of up to two audit servers. The configuration is done by issuing the **audit_config_set** XCLI command, as shown in Example 5-9. This command configures a primary and an optional secondary auditing server for command logging.

Example 5-9 Configuring the audit servers

```
audit_config_set primary_server=Address [ primary_port=port ] [
secondary_server=Address ]
```

The command will not check whether the specified server can be pinged or is listening on the specified port.

Note: This feature is domain-unaware, it is not possible to define auditing at the domain level.

Next, to effectively enable auditing, you must issue an **audit_enable** command. At least the primary_server must be configured for **audit_enable** to succeed.

Enabling or disabling of this feature is restricted to the security administrator.

The following new XCLI commands were introduced with XIV release 11.5:

- **audit_enable**

  To effectively enable auditing. For this command to complete successfully, at least one syslog server must be configured

- **audit_disable**

  A prerequisite for this is that auditing is currently enabled (displayed as “yes” in **audit_show**)

- **audit_show**

  Results of this command indicate whether auditing is currently enabled or disabled.

- **audit_config_set**

  This command is used to configure the parameters required for enabling audits. Currently, the only supported protocol is SYSLOG (over UDP).

- **audit_config_get**

  This command displays the current audit-related configuration.

  Example 5-10 shows a single-server configuration.

  Example 5-10

```
XIV_PFE2_1340010>>audit_config_get
Primary Server   Primary Port   Secondary Server   Secondary Port   Protocol
10.0.0.222       514                               0                syslog
```

5.6.2 Define Idle Timeout to lock a GUI or XCLI session

Users can now specify a session timeout to indicate the idle time for closing the GUI and XCLI session. This feature enables clients to comply with their companies’ security standards and policies. It is also required by the PCI-DSS.

The XCLI and GUI will be locked after the specified period of idle time, and require reauthentication. By default, the inactivity time interval is set to 15 minutes.

To change the idle timeout parameter, navigate to menu bar on upper left of the main window then select **Management** from the **Tools** drop down menu as shown in Figure 5-45 on page 244.

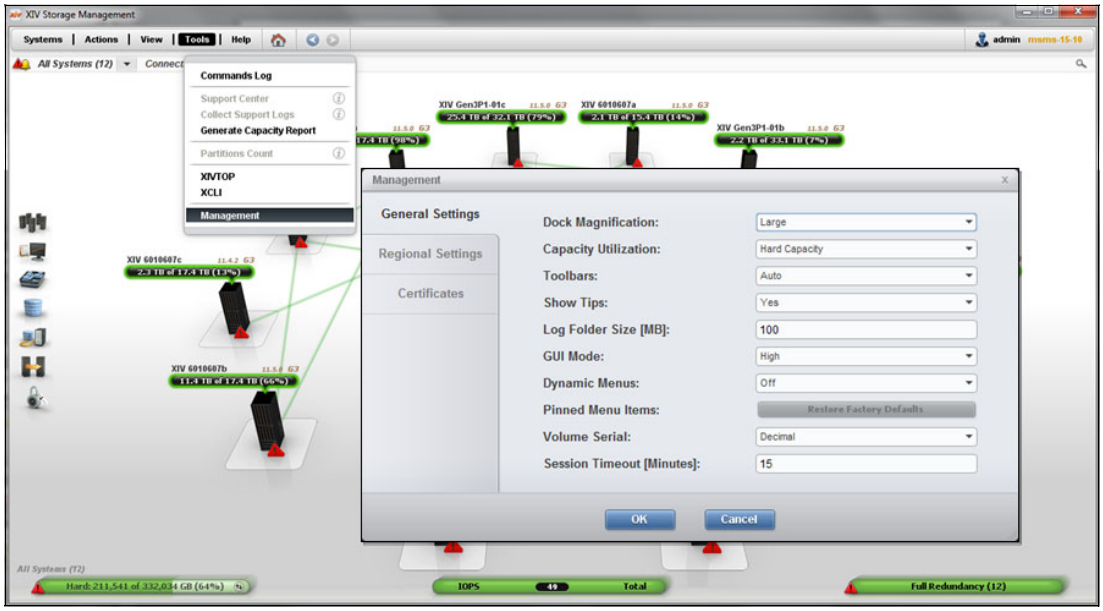


Figure 5-45 Changing idle timeout for lock GUI/XCLI session

5.7 LDAP-based authentication

The XIV Storage System offers the capability to use LDAP server-based user authentication.

When LDAP authentication is enabled, the XIV Storage System accesses a specified LDAP directory to authenticate users whose credentials are maintained in the LDAP directory (except for the admin, technician, maintenance, and development users, which remain locally administered and maintained).

The benefits of an LDAP-based centralized user management can be substantial when considering the size and complexity of the overall IT environment. Maintaining local user credentials repositories is relatively straightforward and convenient when dealing with only a few users and a few storage systems. However, as the number of users and interconnected systems grows, the complexity of user account management rapidly increases and managing such an environment becomes a time-consuming task.

In this section, we review various benefits of this approach. Although the benefits from using LDAP are significant, you must also evaluate the considerable planning effort and complexity of deploying LDAP infrastructure, if it is not already in place.

5.7.1 Introduction to Lightweight Directory Access Protocol

The Lightweight Directory Access Protocol (LDAP) is an open industry standard that defines a standard method for accessing and updating information in a directory.

A directory is a listing of information about objects arranged in an order that gives details about each object. Common examples are a city telephone directory and a library card catalog. In computer terms, a directory is a specialized database, also called a *data repository*, that stores typed and ordered information about objects. A particular directory might list information about users (the objects) consisting of typed information, such as user names, passwords, and email addresses. Directories allow users or applications to find resources that have the characteristics needed for a particular task.

Directories in LDAP are accessed by using the client/server model. An application that wants to read or write information in a directory does not access the directory directly, but uses a set of programs or application programming interfaces (APIs) that cause a message to be sent from LDAP client to LDAP server. An LDAP server retrieves the information requested on behalf of the client application and returns the requested information if the client has permission to see the information. LDAP defines a message protocol used between the LDAP clients and the LDAP directory servers. This protocol includes methods to search for information, read information, and to update information based on permissions.

5.7.2 LDAP directory components

An LDAP directory is a collection of objects organized in a tree structure. The LDAP naming model defines how objects are identified and organized. Objects are organized in a tree-like structure called the *Directory Information Tree* (*DIT*). Objects are arranged within the DIT based on their distinguished name (DN). Distinguished name defines the location of an object within the DIT. Each object is also referred to as an entry in a directory belonging to an object class. An object class describes the content and purpose of the object. It also contains a list of attributes, such as a telephone number or surname, that can be defined in an object of that object class.

As shown in Figure 5-46, the object with the DN `cn=mbarlen, ou=Marketing, o=IBM` belongs to object class `objectClass=ePerson`.

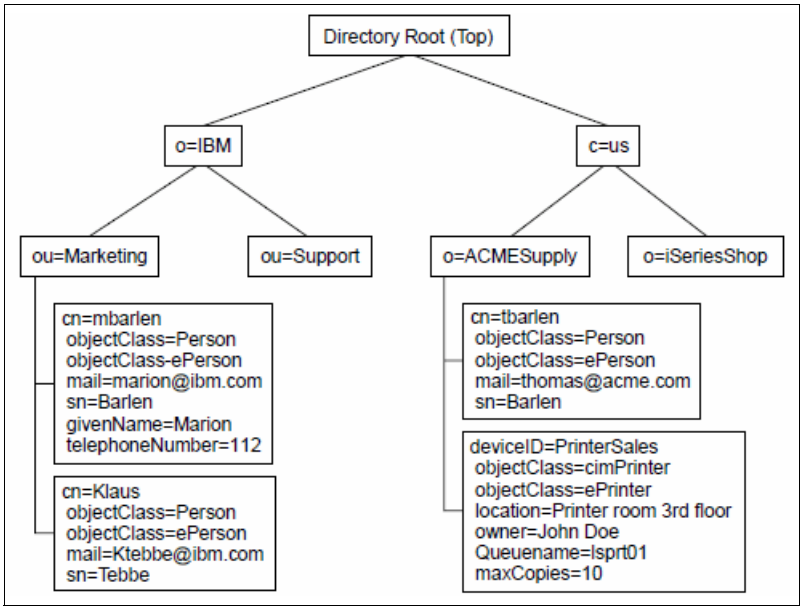


Figure 5-46 Example of a Directory Information Tree (DIT)

Object class ePerson contains the following attributes:

- cn (common name)
- mail
- sn (surname)
- givenName
- telephoneNumber

Each attribute has a value assigned to it:

- cn=mbarlen
- mail=marion@ibm.com
- sn=Barlen
- givenName=Marion
- telephoneNumber=112

In this example, the object represents a single employee record. If a record for a new employee in organizational unit (ou), Marketing, of organization (o), IBM, needs to be created, the same location in DIT is the same, ou=Marketing, o=IBM. Additionally, the same set of attributes defined by objectClass `ePerson` are also used. The new object is defined using its own set of attribute values because the new employee will have their own name, email address, phone number, and so on.

For more information about the directory components, see *Understanding LDAP - Design and Implementation*, SG24-4986.

All the objects and attributes with their characteristics are defined in a schema. The schema specifies what can be stored in the directory.

5.7.3 LDAP product selection

LDAP authentication for Version 10.2.x or 11.x of the XIV Storage System supports three LDAP server products:

- Microsoft Active Directory
- Oracle Directory Server Enterprise Edition
- OpenLDAP

The current skill set of your IT staff is always an important consideration when choosing any products for centralized user authentication. If you have skills in running a particular directory server, it might be a wise choice to standardize on this server because your skilled people will best be able to customize and tune the server. Your experts will be able to provide the most reliable and highly available implementation for the LDAP infrastructure.

Microsoft Active Directory might be a better choice for an enterprise with most of its infrastructure components deployed using Microsoft Windows operating system. Oracle Java Systems Directory and Oracle Directory Server Enterprise Edition provide support for UNIX-like operating systems (including Linux), and Microsoft Windows. OpenLDAP is simple, small, and easy to set up.

All LDAP servers share many basic characteristics because they are based on the industry standards Request for Comments (RFC). However, because of implementation differences, they are not always entirely compatible with each other. For more information about RFCs, particularly regarding LDAP RFC 4510-4533, see the following website:

`http://www.ietf.org/rfc.html`

Current implementation of LDAP-based user authentication for XIV Storage System does not support connectivity to multiple LDAP servers of various types. However, it is possible to configure an XIV Storage System to use multiple LDAP servers of the same type to eliminate a single point of failure. The XIV Storage System supports communication with only one LDAP server at a time. The LDAP authentication configuration allows specification of multiple LDAP servers that the XIV Storage System can connect to if a specified LDAP server is inaccessible.

5.7.4 LDAP role mapping

Before any LDAP user can be granted access to XIV Storage System, the user must be a member of an appropriate LDAP group.

There are three predefined role mappings when using LDAP:

storageadmin	Defined as a single LDAP group.
securityadmin	Defined as a single LDAP group.
readonly	Defined as a single LDAP group.
applicationadmin	Defined as a single LDAP group per XIV User Group.

Important: An LDAP user cannot be a member of more than one LDAP group, so it cannot be associated with more than one XIV Storage System role mapping.

In native mode, a role is explicitly assigned to a user at the time of *user account* creation. In LDAP mode, the role of a specific user is determined at the time the user logs in to an XIV Storage System.

Planning considerations

When initially planning to use LDAP-based authentication with XIV Storage System, the LDAP server administrator must decide on which LDAP attribute can be used for role mapping. As described in 5.7.2, “LDAP directory components” on page 245, each LDAP object has several associated attributes. The type of LDAP object classes used to create a user account for XIV Storage System authentication depends on the type of LDAP server being used. The Oracle Directory server and OpenLDAP use the `inetOrgPerson` LDAP object class, and Active Directory uses the `organizationalPerson` LDAP object class for definition of user accounts for XIV Storage System authentication.

For a definition of the `inetOrgPerson` LDAP object class and list of attributes, see the Internet FAQ archive website:

`http://www.faqs.org/rfcs/rfc2798.html`

For a definition of the `organizationalPerson` LDAP object class and list of attributes, see the Microsoft website:

`http://msdn.microsoft.com/en-us/library/ms683883(VS.85).aspx`

In our illustration, we use both the Active Directory `memberOf` attribute and the `OpenLDAP` description attribute for role mapping.

The mapping is done by assigning the appropriate attribute value to the **`xiv_group_attrib`** configuration parameter with the **`ldap_config_set`** XCLI command:

```
ldap_config_set xiv_group_attrib=memberOf
```

It can also be defined in the XIV Storage Management GUI, as described in the next section, “LDAP role mapping for the storageadmin and readonly roles”.

LDAP role mapping for the storageadmin and readonly roles

The XIV Storage System administrator requires two LDAP group names that are used for storageadmin and readonly role names in LDAP.

In our example, the XIV Storage System administrator uses the “XIVAdmins” and “XIVReadonly” LDAP group names for mapping to storageadmin and readonly roles. This

mapping needs to be entered using the full DN because this full DN is the populated value in the `memberOf` attribute.

The XIV Storage System administrator sets corresponding parameters in the XIV Storage System:

```
ldap_config_set
storage_admin_role="CN=XIVAdmins,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com"
ldap_config_set
read_only_role="CN=XIVReadonly,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com"
```

Case sensitivity: The LDAP server does not use case-sensitive string matching for the `memberOf` attribute value. For example, `XIVAdmins`, `xivadmins`, or `XIVADMINS` are recognized as equal strings. To simplify XIV Storage System administration, however, treat both the XIV Storage System configuration parameter and LDAP attribute value as though they were case-sensitive and assign the `XIVAdmins` value to both.

The `XIVAdmins` and `XIVReadonly` names are used because both strings can be easily associated with their corresponding XIV Storage System roles: storageadmin and readonly. It is not necessary to use the same names in your XIV Storage System configuration.

However, if you were to change these parameters, consider using names that are self-descriptive and easy to remember, to simplify the LDAP server administration tasks. Every time the LDAP server administrator creates a new XIV System Storage account, one of the names must be entered as a description attribute value (except for the applicationadmin role, as we explain in “LDAP role mapping for the applicationadmin role” on page 250). After being configured in both XIV and LDAP, changing these parameters, although possible, can potentially be time consuming because each existing LDAP account must be changed individually to reflect the new attribute value.

The configuration tasks can also be done from the XIV Storage Management GUI. On the main XIV Storage Management window, click **Systems** → **System Settings**, select the **LDAP** option, and select the **Role Mapping** tab. Enter the following variables, as shown in Figure 5-47 on page 249:

- In the XIV Group Attribute field:

  `memberOf`

- In the Storage Admin Role field:

  `CN=XIVAdmins, CN=Users,DC=itso,DC=storage,DC=ibm,DC=com`

- In the Read Only Role field:

  `CN=XIVReadonly,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com`

Click **Update** to save your changes.

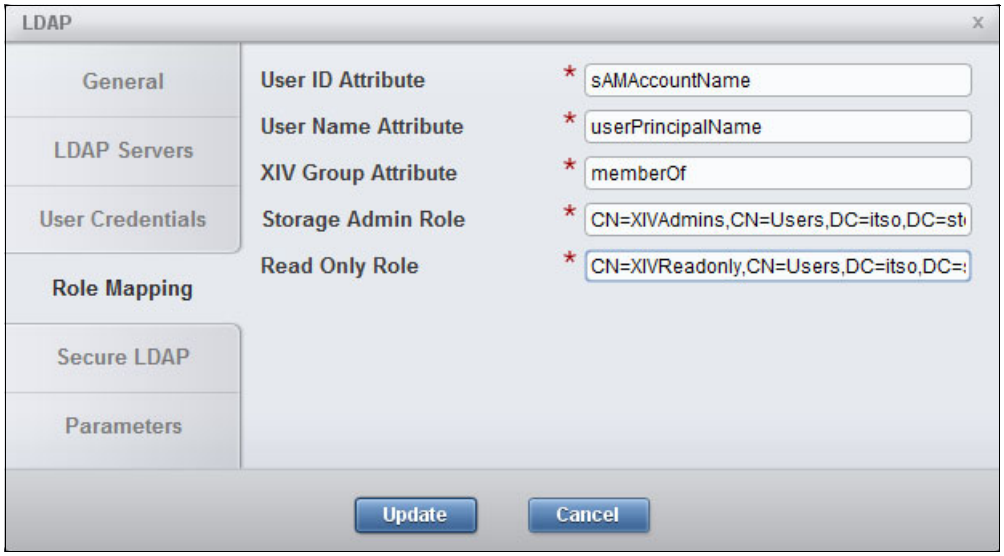


Figure 5-47 LDAP definitions in XIV Storage Management GUI

There is also an XIV Storage Management GUI wizard available to help you with configuring LDAP and role mappings. See 5.8, “Defining LDAP on the XIV Storage System” on page 251.

The XIV Storage System administrator informs the LDAP administrator that the `XIVAdmin` and `XIVReadonly` groups must be used for role mapping.

Tip: When using role mapping using group memberships (such as using the `memberOf` attribute with Active Directory), you must specify the full DN string of the LDAP group name.

The LDAP server returns the full DN string back to the XIV Storage System during the authentication response and it must match the same path.

The LDAP administrator creates a user account in LDAP and assigns the user to the `memberOf` attribute. When the newly created user logs in to the system, XIV systems perform the role mapping, as shown in Figure 5-48 on page 250.

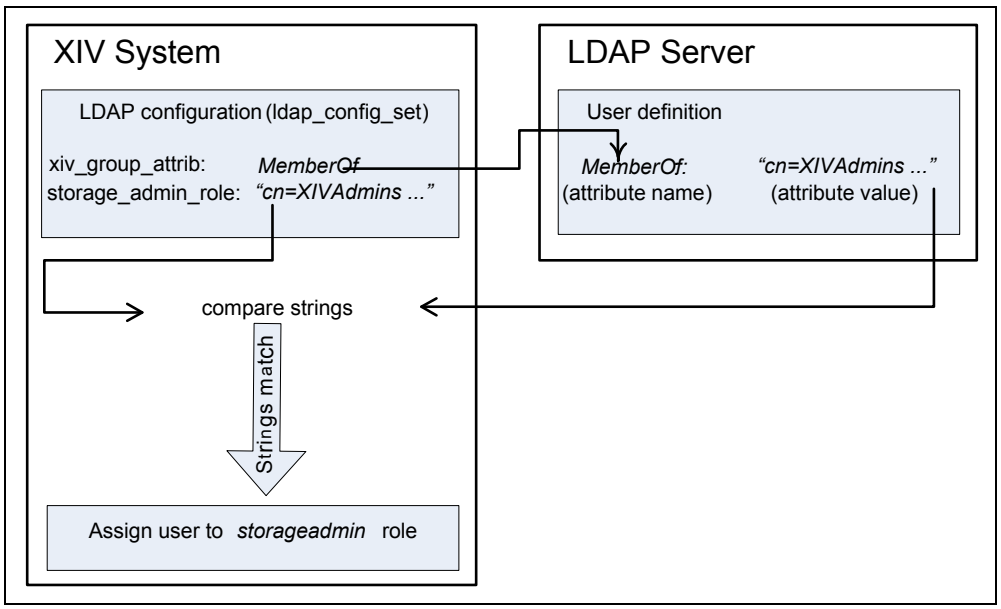


Figure 5-48 Assigning an LDAP authenticated user to the storageadmin role

LDAP role mapping for the applicationadmin role

The LDAP account can also be assigned to an applicationadmin role, but the mechanism of creating role mapping in this case differs from the one used for storageadmin and readonly role mapping.

The XIV Storage System assigns a user to the applicationadmin role if it can match the value of the `MemberOf` attribute with the **`ldap_role`** parameter of any user groups defined in the XIV Storage System. If an account is assigned the applicationadmin role, it also becomes a member of the user group whose **`ldap_role`** parameter matches the value of the user's `MemberOf` attribute.

The user group must be created before the user logs in to the system or the login fails. The XIV Storage System administrator creates a user group, using the **`user_group_create`** XCLI command as follows:

```
user_group_create user_group=app01_group ldap_role="cn=app01_admins
,cn=Users,dc=itso,dc=storage,dn=ibm,dc=com"
```

After the LDAP administrator creates the user account and assigns the `app01_administrator` value to the `MemberOf` attribute, the user can be authenticated by the XIV Storage System. The role assignment and group membership inheritance for a newly created user is shown in Figure 5-49.

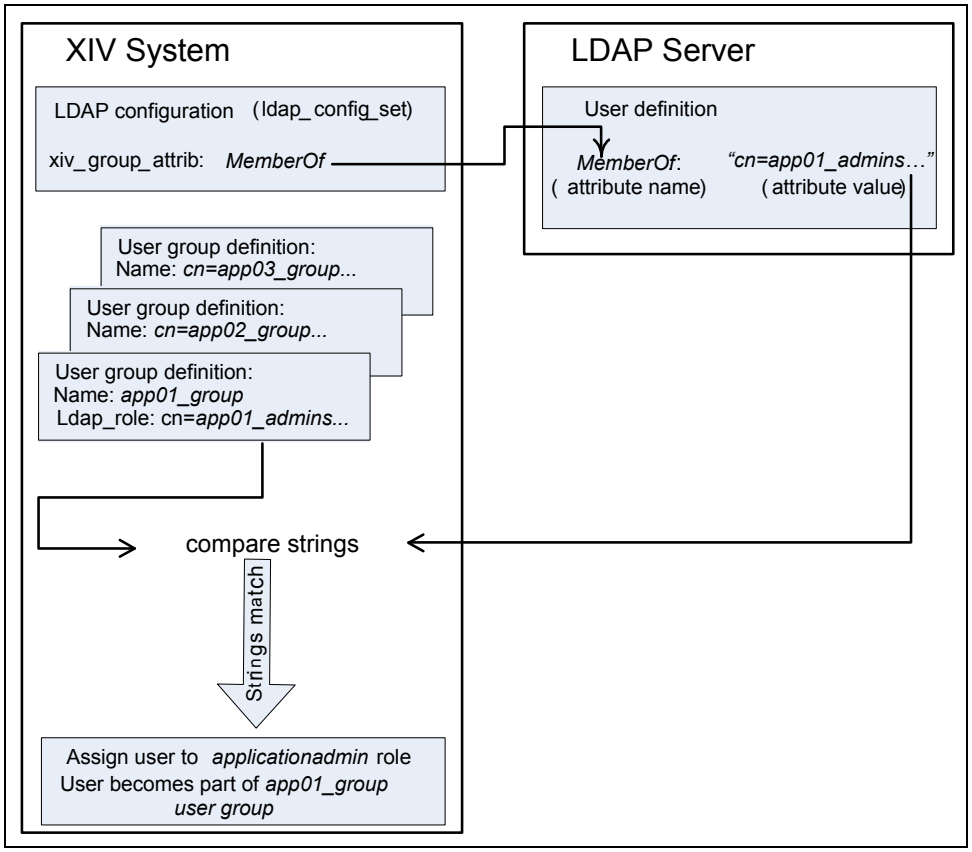


Figure 5-49 Assigning an LDAP authenticated user to the applicationadmin role

If the XIV Storage System cannot find a match for the value assigned to the `MemberOf` attribute of a user, the user is denied system access.

5.8 Defining LDAP on the XIV Storage System

This section describes how to configure the XIV Storage System in an existing LDAP environment.

Important: When upgrading the XIV Storage System Software (firmware), any previous LDAP configuration on the XIV Storage System might be reset. Ensure that you have saved the LDAP configuration information if you need to reload it.

All users (but not the predefined users) that you eventually created for native authentication are deactivated (not deleted) when you activate the LDAP mode in the XIV Storage System. Defined users no longer appear in the XIV Storage System user settings. However, they are not deleted and are shown and activated again if you eventually disable the LDAP authentication.

There are three ways to configure the LDAP on the XIV Storage System:

- The XIV Storage Management GUI LDAP wizard
- The regular XIV Storage Management GUI windows
- The XCLI

You can see examples for all methods in the following sections.

The first step for the XIV Storage System administrator to do is to verify domain name server (DNS) name resolution, as shown in Example 5-11.

Example 5-11 DNS name resolution verification

```
>> dns_test name=itso.storage.ibm.com
Name                    IP (Primary DNS)   IP (Secondary DNS)
itso.storage.ibm.com    9.155.113.143
```

If the **dns_test** command returns an unexpected result, do not proceed further with the configuration steps until the DNS name resolution issue is resolved.

Important: As a preferred practice, the LDAP server and XIV Storage System must have their clocks synchronized to the same time source, be registered, and be configured to use the same DNS servers.

5.8.1 Using the XIV Storage Management GUI LDAP wizard to configure LDAP

To configure LDAP by using the XIV Storage Management GUI LDAP wizard, complete the following steps:

1. Right-click the Access icon and click **Users**, as shown in Figure 5-50.

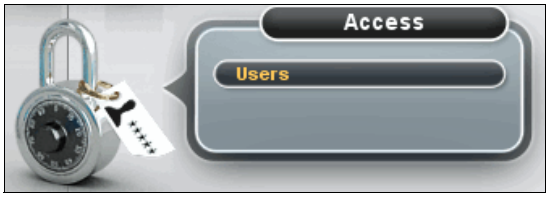


Figure 5-50 Users

2. Click the **LDAP Wizard** icon on the toolbar, as shown in Figure 5-51.

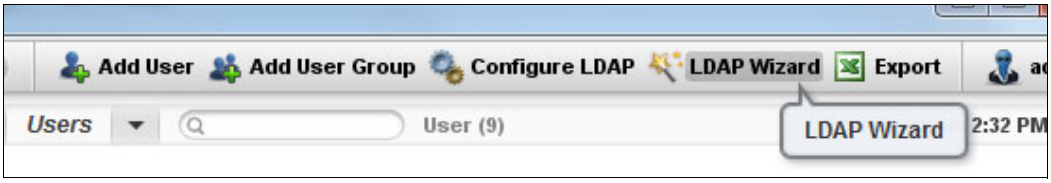


Figure 5-51 LDAP Wizard icon

The LDAP wizard opens the LDAP Configuration Welcome window. Click **Next** to start the configuration.

In the Server Type tab, as shown in Figure 5-52 on page 253, you can choose the Server Type from the corresponding drop-down menu. You can use either a Microsoft Active Directory or an Oracle Directory Server.

3. Click **Next** to choose Microsoft Active Directory, as we did in this example.

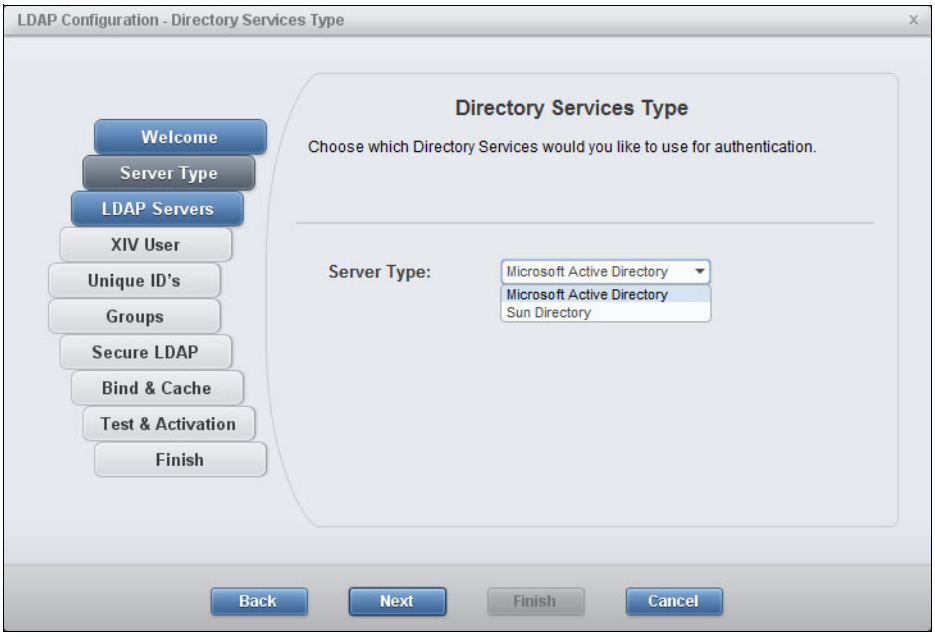


Figure 5-52 Server Type tab

4. From the LDAP Servers tab, click the plus sign (**+**) icon at the right to add an LDAP server to the configuration, as shown in Figure 5-53.

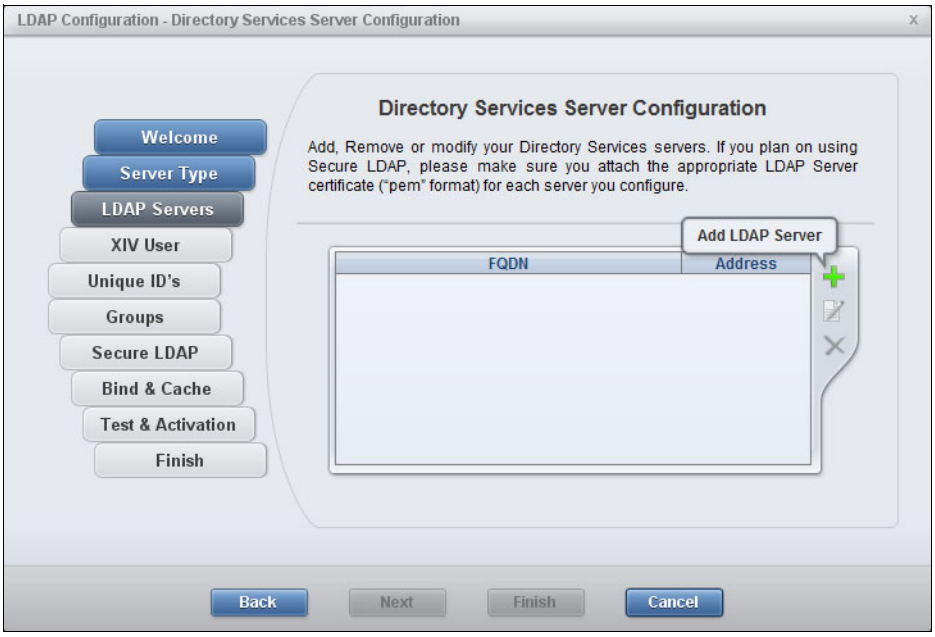


Figure 5-53 Adding LDAP servers

5. In the window that opens (Figure 5-54), enter the fully qualified domain name (FQDN), which is the DNS name of the LDAP server (such as `itso.storage.ibm.com` in our example).

   Also, enter the LDAP server IP address and the Search DN, which in our Active Directory implementation is CN=Users,DC=itso,DC=storage,DC=ibm,DC=com, as shown in Figure 5-54.

   For example, in a default Active Directory implementation, a server with the domain name `ldap.domain.de` maps to CN=Users,DC=ldap,DC=domain,DC=de.

6. If you need a certification file for the connection to your LDAP server, add it here for the connection.
7. If your LDAP server uses nonstandard port numbers, enter those accordingly. Leave these fields blank to use the standard ports (server port 389, secure server port 636).
8. Click **Create** to save the settings.

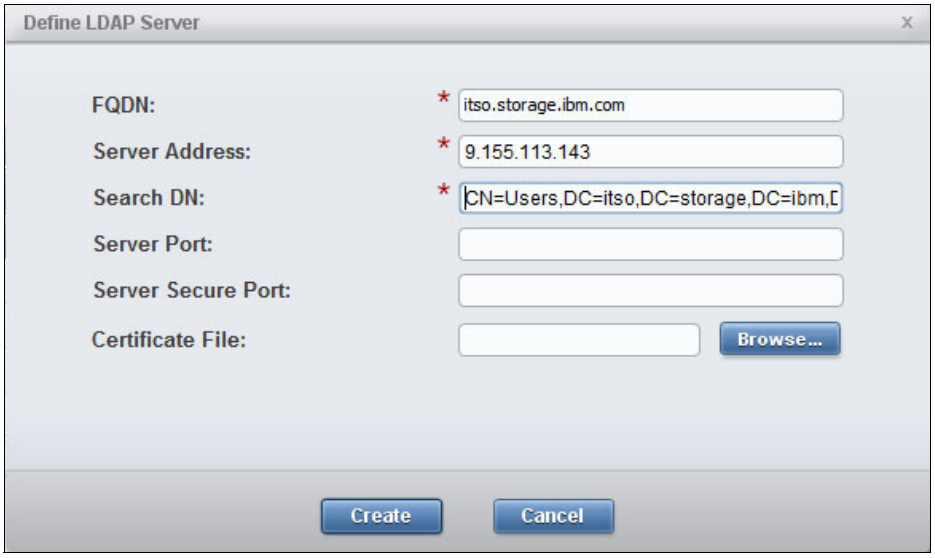


Figure 5-54 Define LDAP Server window

Now you can see your added server reflected in the table, as shown in Figure 5-55.

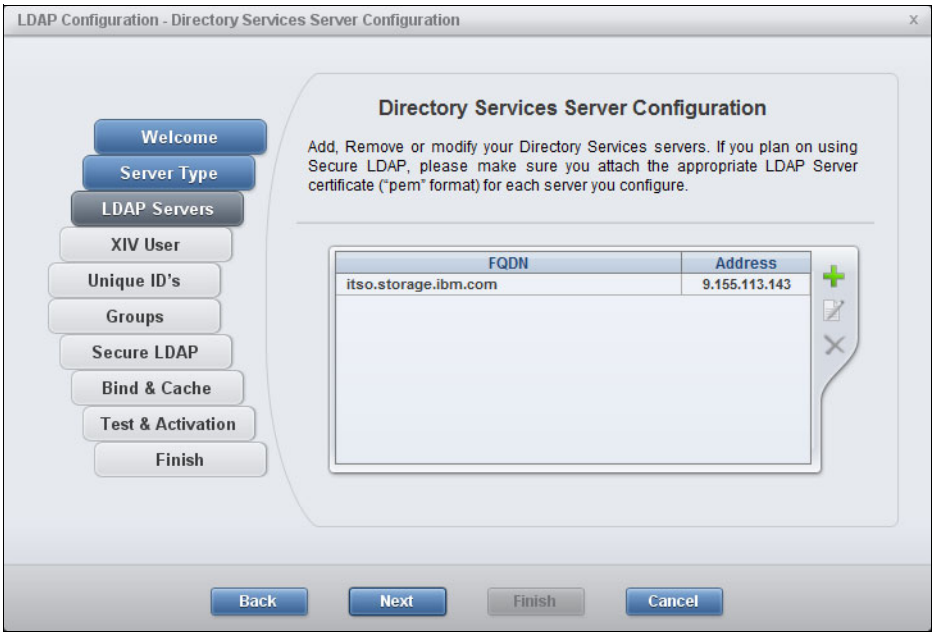


Figure 5-55 LDAP Servers tab

9. Add more LDAP servers, if necessary, or click **Next** to continue with the configuration process.
10. From the XIV User tab, as shown in Figure 5-56, define one user (this user must be an already defined LDAP user) that will be used by XIV to validate the LDAP settings. That user does not need any special permissions.
11. As shown in Figure 5-56, enter the complete user DN for LDAP, which in our example is CN=XIV,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com. Enter the password for this user as well. The password is not saved; it is only used for this wizard process. Click **Next**.

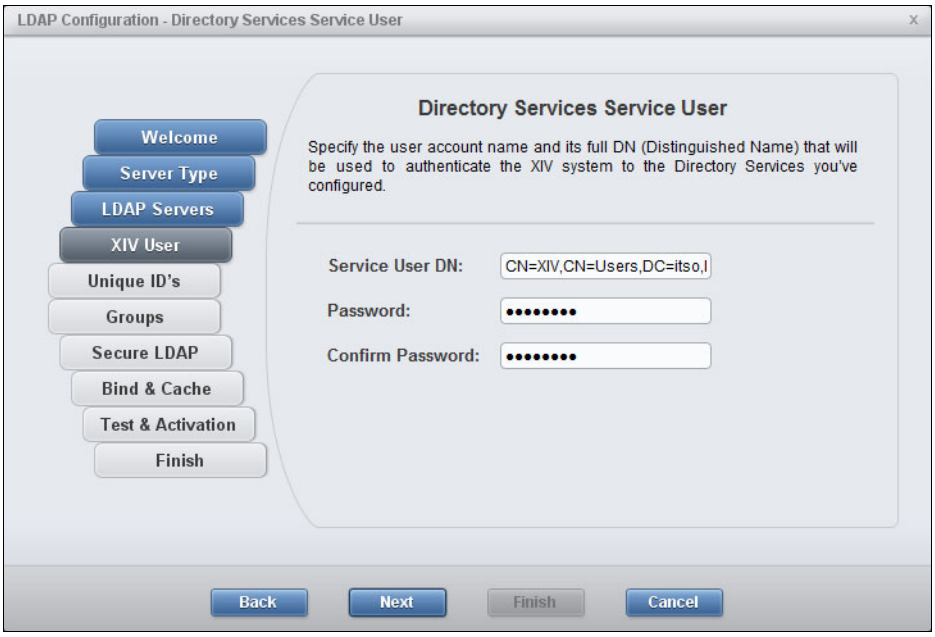


Figure 5-56 XIV User tab

12. In the Unique ID's tab, set the parameters for the login (see Figure 5-57). In our example, the default selections are used. The ID Attribute is set to `sAMAccountName`, the Name Attribute to `userPrincipalName`, and the Group Attribute to `memberOf`. In Oracle Directory Server environments, for the Name Attribute, use `uid`, and for Group Attribute in the XIV Storage System, use `isMemberOf`.

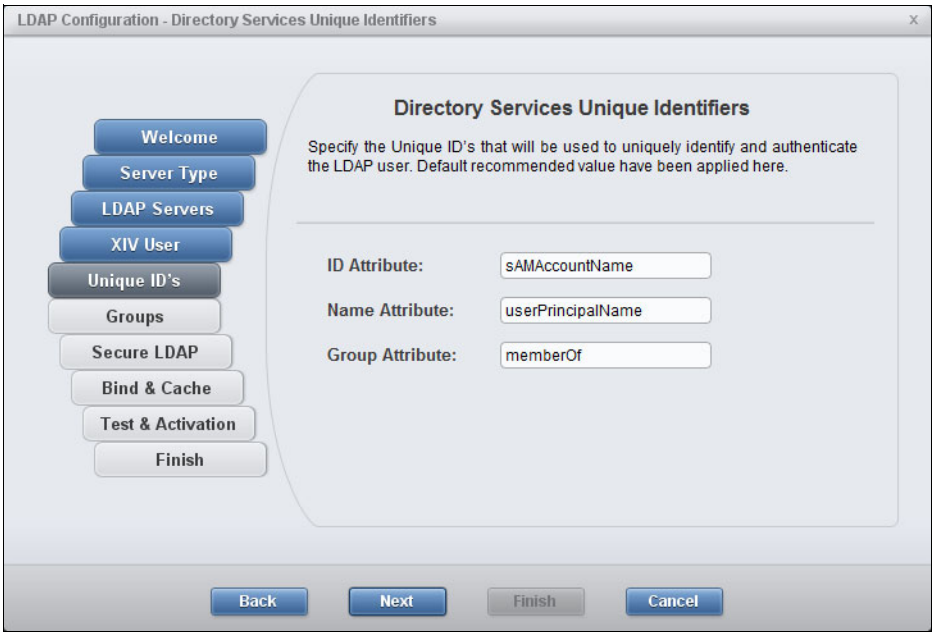


Figure 5-57 Unique ID's tab

13. Ask your LDAP admin for the value of those parameters in your environment. Click **Next** to save the settings and proceed.
14. In the Groups tab that is shown in Figure 5-58 on page 257, you must define the group attributes that the XIV Storage System considers for the permissions. These groups must be defined in LDAP.

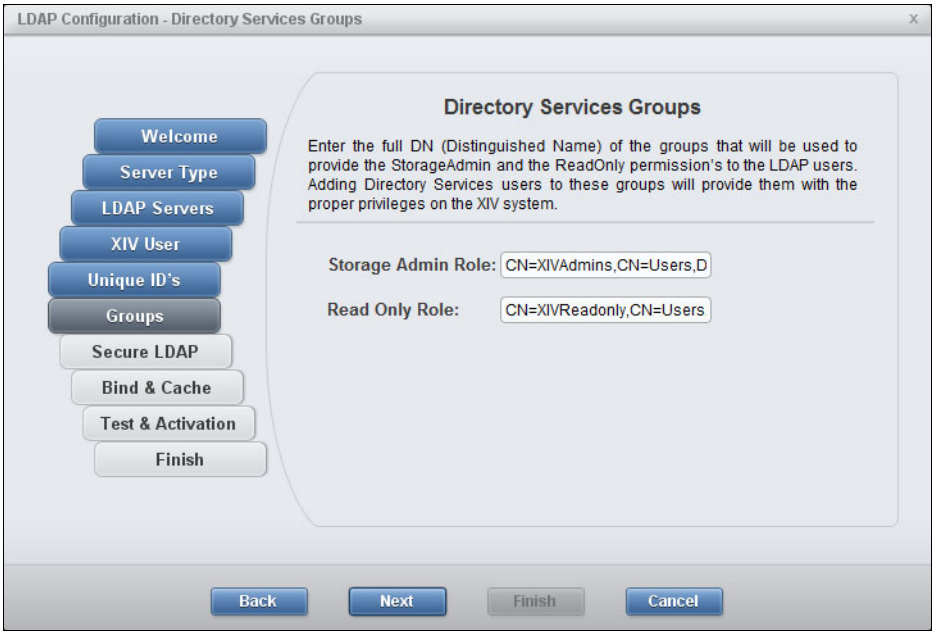

LDAP Configuration - Directory Services Groups
Welcome
Server Type
LDAP Servers
XIV User
Unique ID's
Groups
Secure LDAP
Bind & Cache
Test & Activation
Finish
Directory Services Groups
Enter the full DN (Distinguished Name) of the groups that will be used to provide the StorageAdmin and the ReadOnly permission's to the LDAP users. Adding Directory Services users to these groups will provide them with the proper privileges on the XIV system.
Storage Admin Role:
CN=XIVAdmins,CN=Users,D
Read Only Role:
CN=XIVReadonly,CN=Users
Back
Next
Finish
Cancel

Figure 5-58 Groups tab

Our example has an LDAP group named `XIVAdmins` and another one named `XIVReadonly`.

15. Define the role with the complete DN path. In our example, the Admin role is CN=XIVAdmins,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com and the Read Only role is CN=XIVReadonly,CN=Users,DC=ibm,DC=storage,DC=ibm,DC=com, as shown in Figure 5-58. Click **Next** until you reach the Test and Activation tab.
16. From the Test and Activation tab that is shown in Figure 5-59, enable LDAP by setting Use LDAP to **Yes**. You can also test your LDAP with a user already defined in LDAP.

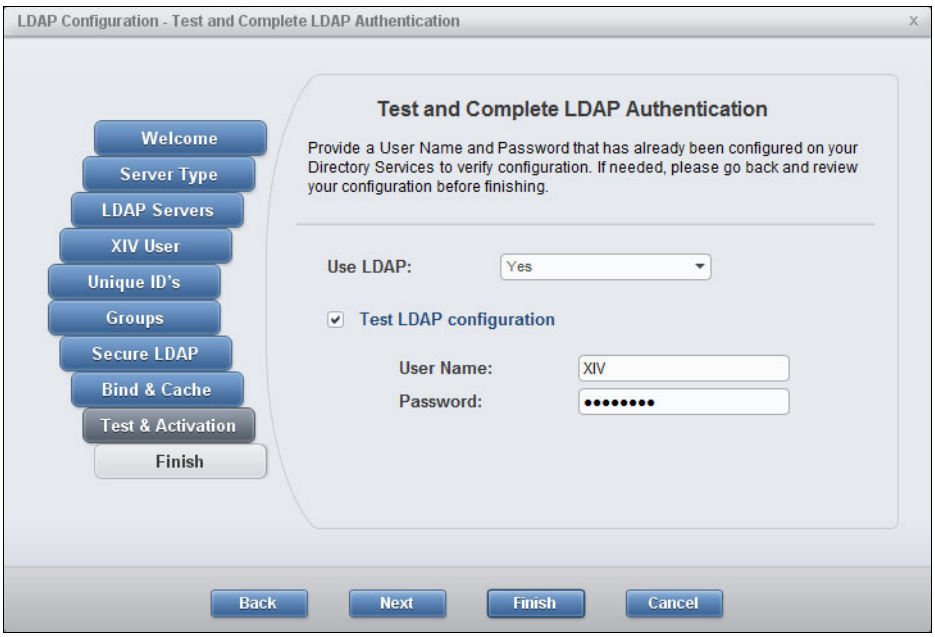


Figure 5-59 Test and Activation tab

17. At this stage, click **Finish** to complete the setup, or click **Next** to go to the exit and summary window that is shown in Figure 5-60.

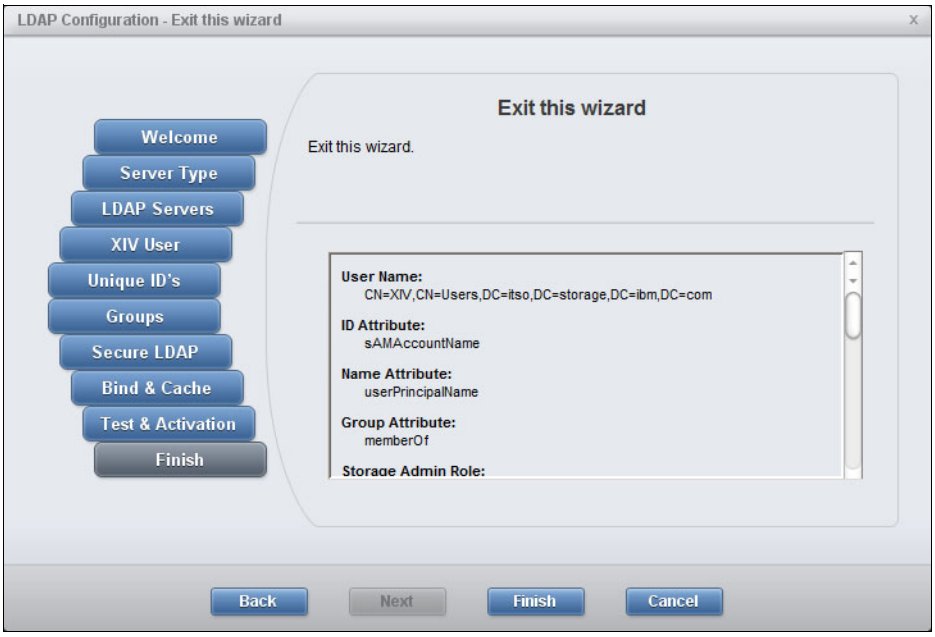


Figure 5-60 Wizard summary

LDAP mode considerations:

- After you activate the LDAP mode, it deactivates all previously defined XIV Storage System users. Only the LDAP user with the appropriate group role has access to the system. You will also be logged out from the XIV Storage Management GUI.
- All local users and passwords will be reactivated again after you have disabled the LDAP mode and return to the native mode. This situation means that the users are deactivated and not deleted.

18. Check the summary and click **Finish**. You might be asked if you are sure that you want to enable LDAP now because it will log you off the system afterward. See Figure 5-61.

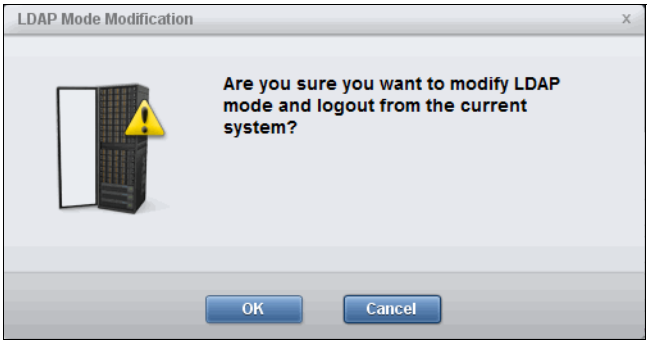


Figure 5-61 Activation confirmation

19. Click **OK**.

20. When you get a confirmation that LDAP was tested and configured, as shown in Figure 5-62, click **OK** again to close the window.

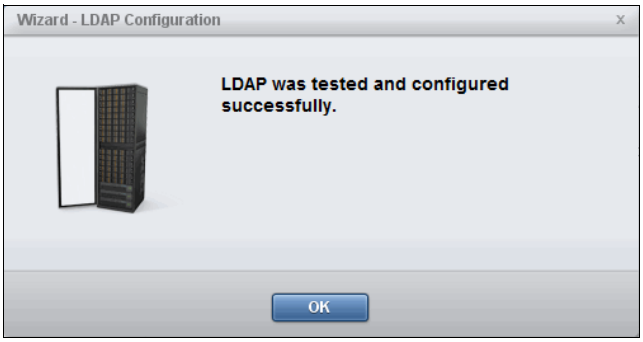


Figure 5-62 End of LDAP configuration

5.8.2 Using the XIV Storage Management GUI directly to configure LDAP

Configuring LDAP in an XIV Storage System can also be done directly through the XIV Storage Management GUI, rather than using the LDAP wizard. Complete the following steps:

1. Hover the cursor over the Access icon and click **Users**, as shown in Figure 5-63.

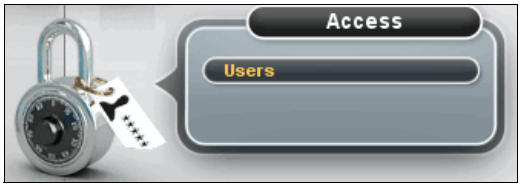


Figure 5-63 Users tab

2. Click the **Configure LDAP** icon on the toolbar, as shown in Figure 5-64.

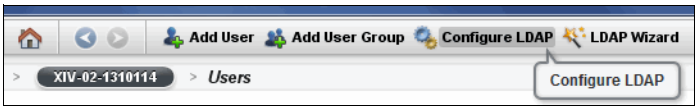


Figure 5-64 Configure LDAP icon

3. In the LDAP configuration menu that is shown in Figure 5-65 on page 260, set the value of Use LDAP to **Yes** and choose the LDAP Server Type that you want to use. In our example, we selected **Microsoft Active Directory**.

   Then, click the **LDAP Servers** tab to go to the next step.

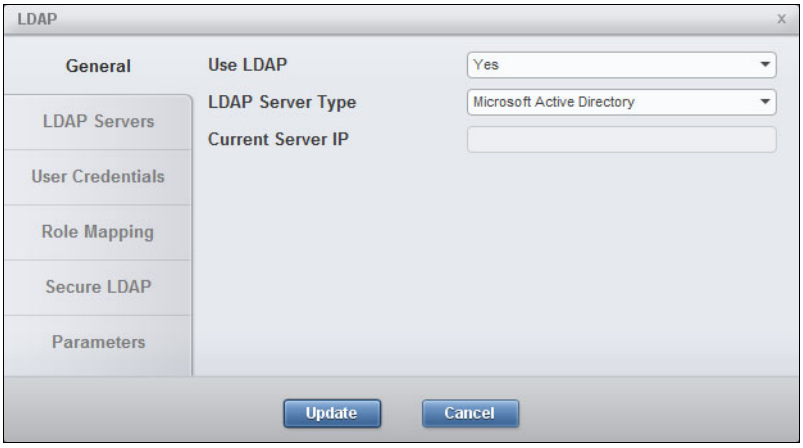


Figure 5-65 LDAP configuration: General

4. In the LDAP Servers window in the LDAP configuration, click the green plus (**+**) icon, as shown in Figure 5-66.

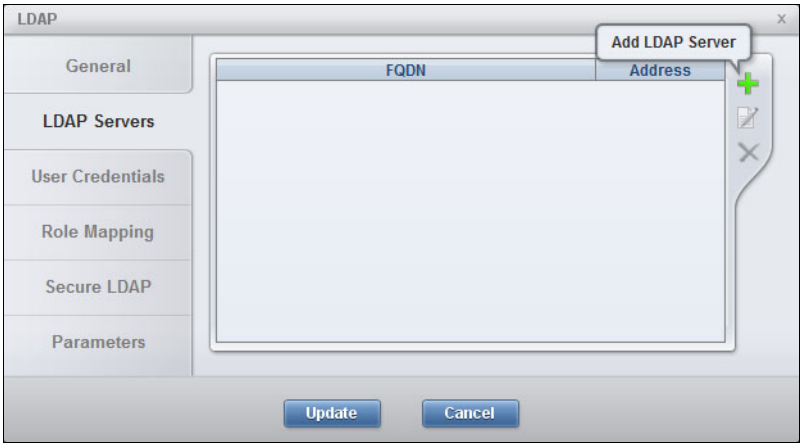


Figure 5-66 LDAP configuration: Servers

5. In the window that opens, enter the FQDN, which is the DNS name of the LDAP server (such as `itso.storage.ibm.com` in our example).

   Also, enter the LDAP server IP address and the Search DN, which in our Active Directory implementation is CN=Users,DC=itso,DC=storage,DC=ibm,DC=com, as shown in Figure 5-67 on page 261.

   If your LDAP server uses nonstandard port numbers, enter those accordingly. Leave these fields blank to use the standard ports (server port 389, secure server port 636).

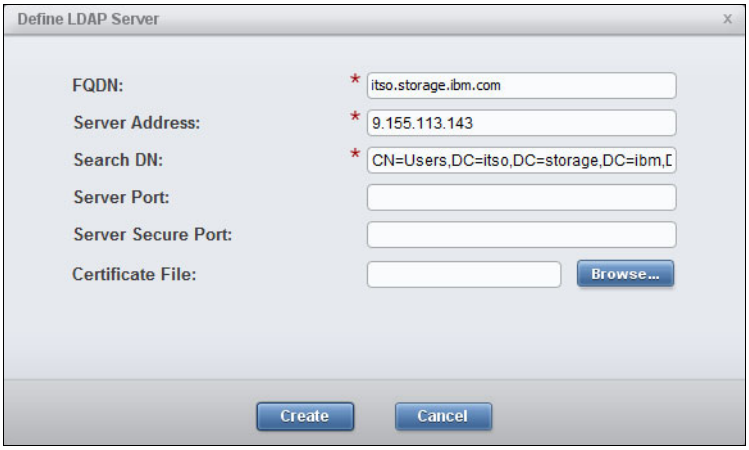


Figure 5-67 Define LDAP servers

In a default Active Directory implementation, for example, a server with the Domain Name `ldap.domain.de` maps to CN=Users,DC=ldap,DC=domain,DC=de.

6. Click **Create** to save these settings.
7. Click **User Credentials** to define the Service User DN, which is used to verify your LDAP settings.

   In Figure 5-68, you can see that a service user named `XIV` is defined using a complete Service User DN, which in this example is CN=XIV,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com.

   Enter and confirm the user password.

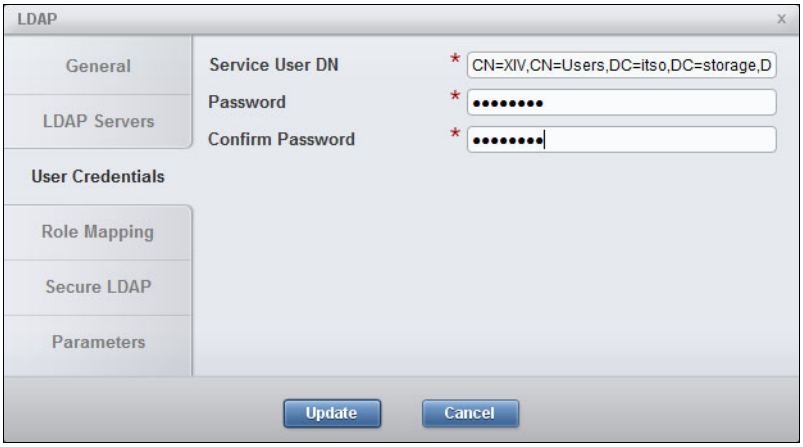


Figure 5-68 LDAP Configuration: User Credentials tab

8. Click **Role Mapping** to enter the LDAP attributes that you want to use for the XIV Storage System login. See Figure 5-69 for an illustration of the values for Microsoft Active Directory that we used in our example:
   - User ID Attribute: sAMAccountName
   - User Name Attribute: displayName
   - XIV Group Attribute: memberOf
   - Storage Admin Role: CN=XIVAdmins,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
   - Read Only role: CN=XIVReadonly,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com

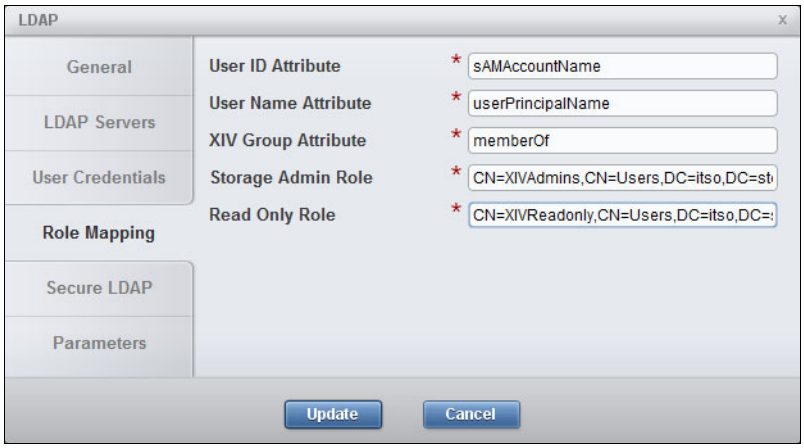


Figure 5-69 LDAP Configuration: Role Mapping tab

LDAP mode: When you activate the LDAP mode, it deactivates all defined users on the XIV Storage System and only LDAP users with the appropriate group role have access to the system. You will also be logged out from the XIV Storage Management GUI.

9. Click **Update** and the configuration is now finished.

5.8.3 Using XIV Storage System command-line interface

The quickest method for configuring LDAP in an XIV Storage System is to use the XIV Storage System Command-Line Interface (XCLI).

You must define your LDAP Server in the XIV Storage System, as shown in Example 5-12.

Example 5-12 XCLI add server

```
XIV-02-1310114>>ldap_add_server fqdn=itso.storage.ibm.com address=9.155.113.143
base_dn="CN=Users,DC=itso,DC=storage,DC=ibm,DC=com"
Command executed successfully.

XIV-02-1310114>>ldap_list_servers
FQDN                  Address         Base DN ...
itso.storage.ibm.com  9.155.113.143   CN=Users,DC=itso,DC=storage,DC=ibm,DC=com.
```

After adding one or more LDAP servers, you can define the login attributes in the XIV Storage System. In Example 5-13, you can see an example for a possible Microsoft Active Directory implementation.

Example 5-13 XCLI LDAP configuration

```
XIV-02-1310114>>ldap_config_set server_type="Microsoft Active Directory"
xiv_group_attrib=memberOf
storage_admin_role=CN=XIVAdmins,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
read_only_role=CN=XIVReadonly,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
xiv_user=CN=XIV,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
user_id_attrib=sAMAccountName user_name_attrib=displayName

Command executed successfully.
```

For a description of the LDAP role mapping, see 5.7.4, “LDAP role mapping” on page 247. The XIV Storage System **storage_admin_role**, **read_only_role** and **xiv_group_attrib** configuration parameters must have values assigned for LDAP role mapping to work.

After the configuration is submitted successfully, you can check your values by running **ldap_config_get**.

Now that all the configuration and verification steps are completed, the LDAP mode can be activated by running **ldap_mode_set**, as shown in Example 5-14.

Example 5-14 Enabling the XIV LDAP mode

```
XIV-02-1310114>>ldap_mode_set mode=Active
Warning:   ARE_YOU_SURE_YOU_WANT_TO_ENABLE_LDAP_AUTHENTICATIONy

Command executed successfully.
```

LDAP mode: When you activate the LDAP mode, it deactivates all defined users on the XIV Storage System, and only LDAP users with the appropriate group role have access to the system.

The LDAP authentication mode is now configured, activated, and ready to be tested. A simple test that can validate the authentication result is to open an XCLI session using the credentials of a newly created Active Directory account `xivtestuser1` and run **ldap_user_test**. This command can be successfully executed only by a user authenticated through LDAP.

5.9 LDAP-managed user authentication

When the XIV Storage System is configured for LDAP authentication, user credentials are stored in the centralized LDAP repository. The LDAP repository is on an LDAP server and is accessed by the XIV Storage System using an LDAP protocol.

5.9.1 LDAP repository credential objects

The LDAP repository maintains the following types of credential objects that are used for XIV Storage System authentication:

- LDAP user name
- LDAP user password
- LDAP user role
- User groups

LDAP user name

The XIV Storage System limitations for acceptable user names, such as number of characters and character set, no longer apply when user names are stored in an LDAP repository. Each LDAP product has its own set of rules and limitations that applies to user names. Generally, do not use long names and non-alphanumeric characters even if your LDAP product of choice supports it. If you eventually decide to migrate user credentials between local and LDAP repositories or vice versa, the task can be greatly simplified if the same set of rules is applied to both local and centralized repositories. In fact, the set of rules enforced by the XIV Storage System for local user names must be used for LDAP as well because it is the most restrictive of the two. For details about XIV Storage System limitations for user names, see “User name” on page 220.

Special consideration must be given to using the “space” character in user names. Although this feature is supported by LDAP, it has a potential for making certain administrative tasks more difficult because the user names in this case must be enclosed in quotation marks to be interpreted correctly.

The same set of locally stored predefined user names exists on the XIV Storage System regardless of the authentication mode. Users *technician* and *admin* are always authenticated locally even on a system with activated LDAP authentication mode. Avoid creating LDAP user accounts with the same names.

If a user account with the same user name is registered in both local and LDAP repositories, and LDAP authentication mode is in effect, LDAP authentication takes precedence, and the XIV Storage System performs authentication using LDAP account credentials. The only exception to this rule is the predefined user names listed in the previous paragraph. To reduce complexity and simplify maintenance, it is generally not desirable to have the same user names registered in local and LDAP repositories.

If a user account was registered in the local repository on the XIV Storage System before the LDAP authentication mode was activated, this account is accessible while LDAP authentication is in effect. The account becomes accessible again upon deactivation of the LDAP authentication mode.

LDAP user passwords

User passwords are stored in the LDAP repository when the XIV Storage System is in LDAP authentication mode. Password management becomes a function of the LDAP server. The XIV Storage System relies entirely on the LDAP server to provide password management functions, such as enforcing initial password resets, password strength, and password expiration. Various LDAP server products provide their own sets of tools and policies for password management.

Figure 5-70 shows a technique that can be used for password management.

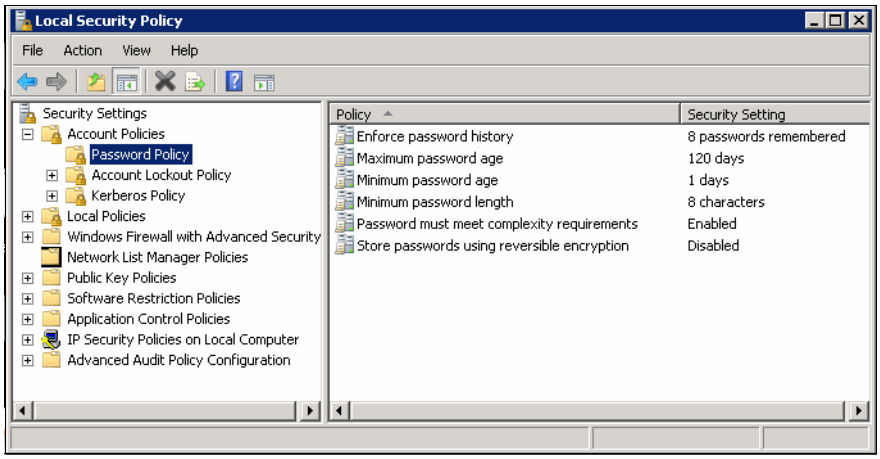


Figure 5-70 Default Active Directory Password Policy settings

If there is a user's password expiration or account lockout, the user receives the message that is shown in Example 5-15 when attempting to log in to XCLI.

Example 5-15 XCLI authentication error because of account lockout

```
>> ldap_test user=xivtestuser1 password=PasswOrd
Error:   LOGIN_FAILURE_USER_NOT_AUTHENTICATED_BY_LDAP_SERVER
Details: User xivtestuser1 was not authenticated by LDAP server
'itso.storage.ibm.com'.
```

The XIV Storage Management GUI in this situation also fails with the error message shown in Figure 5-71.

Figure 5-71 XIV Storage Management GUI authentication failure because of account lockout

Although password policy implementation greatly enhances overall security of the system, all advantages and disadvantages of such implementation must be carefully considered. One of the possible disadvantages is increased management impact for account management as a result of implementing complex password management policies.

Passwords: A comprehensive solution for user password policy implementation is beyond the scope of this book.

LDAP user roles

There are predefined user roles (also referred to as *categories*) used for day to day operation of the XIV Storage System. In the following section, we describe predefined roles, their level of access, and applicable use. The roles are as follows:

- storageadmin

  The storageadmin (Storage Administrator) role is the user role with the highest level of access available on the system. A user assigned to this role can perform changes on any system resource except for maintenance of physical components or changing the status of physical components. The assignment of the storageadmin role to an LDAP user is done through the LDAP role mapping process. For a description, see 5.7.4, “LDAP role mapping” on page 247.

- applicationadmin

  The applicationadmin (Application Administrator) role provides flexible access control over volume snapshots. Users that are assigned to the applicationadmin role can create snapshots of specifically assigned volumes, perform mapping of their own snapshots to an assigned host, and delete their own snapshots. The user group to which an application administrator belongs determines the set of volumes that the application administrator is allowed to manage. If a user group is defined with `access_all=”yes”`, application administrators who are members of that group can manage all volumes on the system. The assignment of the applicationadmin role to an LDAP user is done through the LDAP role mapping process. For a description, see 5.7.4, “LDAP role mapping” on page 247. A description of user group to host association is provided in “User group membership for LDAP users” on page 266.

- readonly

  As the name implies, users assigned to the readonly role can view only system information. A typical use for the readonly role is a user responsible for monitoring system status, system reporting, and message logging, who cannot be allowed to make changes on the system. The assignment of the readonly role to an LDAP user is done through the LDAP role mapping process. For a description, see 5.7.4, “LDAP role mapping” on page 247.

> **Roles:** There is no capability to add new user roles or to modify predefined roles. In LDAP authentication mode, role assignment can be changed by modifying the LDAP attribute (`memberOf` in our example).

LDAP authentication mode implements a user role mechanism as a form of role-based access control (RBAC). Each predefined user role determines the level of system access and associated functions that a user is allowed to use.

> **RBAC:** The XIV Storage System implements RBAC-based authentication and authorization mechanisms.

All user accounts must be assigned to a single user role. Any LDAP user that is assigned to multiple roles will not be authenticated by the XIV Storage System. Deleting role assignment (by removing the description attribute value in the LDAP object) of LDAP users also leads to the inability of the XIV Storage System to authenticate that user.

User group membership for LDAP users

A user group is a group of application administrators who share a set of snapshot management permissions. The permissions are enforced by associating the user groups with hosts or clusters. User groups are defined locally on the XIV Storage System.

User group membership for an LDAP user is established during the login process by matching the designated LDAP attribute value with the **ldap_role** parameter assigned to a user group. A user group is associated with host volumes through access definition. An LDAP user, which is a member of the user group, is permitted to manage snapshots of volumes mapped to the host associated with the user group.

User groups have the following characteristics in LDAP authentication mode:

- Only users assigned to the applicationadmin role can be members of a user group.
- An LDAP user can only be a member of a single user group.
- A maximum of eight user groups can be created.
- In LDAP authentication mode, there is no limit on the number of members in a user group.
- If a user group is defined with `access_all=”yes”`, users assigned to the applicationadmin role who are members of that group can manage all snapshots on the system.
- The **ldap_role** user group parameter can be assigned only a single value.
- The **ldap_role** parameter must be unique across all defined user groups.
- Only users assigned to the storageadmin role can create, modify, and delete user groups.
- Only users assigned to the storageadmin role can modify the **ldap_role** parameter of a user group.

Important: A user group membership can be defined only for users assigned to the applicationadmin role.

Figure 5-72 shows the relationship between LDAP user, LDAP role, XIV Storage System role, user group membership, associated host, mapped volumes, and attached snapshots.

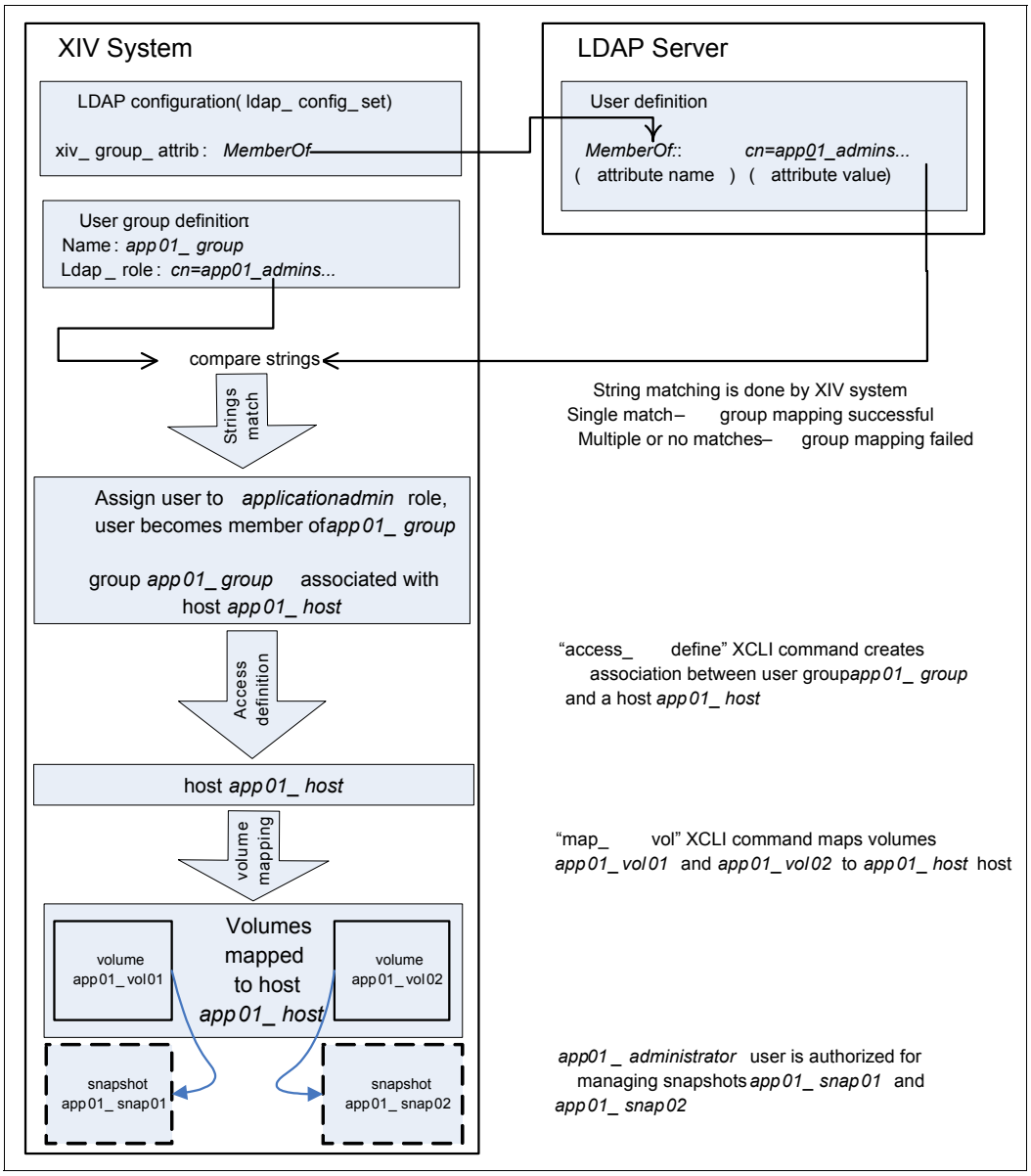

XIV System
LDAP configuration(ldap_ config_ set)
xiv_ group_ attrib : MemberOf
LDAP Server
User definition
MemberOf:: cn=app01_admins...
(attribute name) (attribute value)
User group definition
Name : app01_ group
Ldap_ role : cn=app01_admins...
compare strings
Strings match
String matching is done by XIV system
Single match– group mapping successful
Multiple or no matches– group mapping failed
Assign user to applicationadmin role,
user becomes member of app01_ group
group app01_ group associated with
host app01_ host
Access definition
"access_ define" XCLI command creates
association between user group app01_ group
and a host app01_ host
host app01_ host
volume mapping
"map_ vol" XCLI command maps volumes
app01_ vol01 and app01_ vol02 to app01_ host host
volume app01_ vol01
Volumes mapped to host app01_ host
volume app01_ vol02
snapshot app01_ snap01
snapshot app01_ snap02
app01_ administrator user is authorized for
managing snapshots app01_ snap01 and
app01_ snap02

Figure 5-72 User group membership for LDAP user

5.9.2 Managing LDAP user accounts

Managing user accounts in LDAP authentication mode is done using LDAP management tools. The XCLI commands and XIV Storage Management GUI tools cannot be used for creating, deleting, modifying, or listing LDAP user accounts. The set of tools for LDAP account management is specific to the LDAP server type. The same set of LDAP management tools can also be used for account removal, modification, and listing.

Generating LDAP account lists

To generate a list of all LDAP user accounts registered under the **Base_DN** (the XIV Storage System configuration parameter specifying the location of LDAP accounts in the DIT), use the **ldapsearch** queries that are shown in Example 5-16 and Example 5-17.

Example 5-16 Generating list of LDAP accounts registered in OpenLDAP Directory

```
# ldapsearch -LL -H ldap://localhost:389 -x -b
ou=Users,dc=itso,dc=storage,dc=ibm,dc=com -D
uid=itso,ou=Users,dc=itso,dc=storage,dc=ibm,dc=com -w Passw0rd uid |grep ^uid
uid: itso
uid: zejn
uid: almira
uid: xivtestuser
```

Example 5-17 Generating list of LDAP accounts registered in OpenLDAP Directory

```
# # ldapsearch -x -H ldap://9.155.113.143:389 -x -b
cn=Users,dc=itso,dc=storage,dc=ibm,dc=com -D
cn=itso,cn=Users,dc=itso,dc=storage,dc=ibm,dc=com -w Passw0rd cn |grep ^cn
cn: Users
cn: krbtgt
cn: itso_appuser1
cn: Domain Computers
cn: Domain Controllers
cn: xivtestuser3
...
```

The queries generating LDAP account lists are provided as a demonstration of LDAP tools capabilities to perform a search of information stored in the LDAP directory and generate simple reports. Both queries are issued on behalf of the LDAP administrator account cn=Manager (OpenLDAP), cn="Directory Manager" (Oracle Java Directory), and cn="Administrator" (Active Directory). A privileged account, such as LDAP administrator, has the authority level that allows that person to list, create, modify, and remove other user accounts.

The Active Directory management interface allows you to build custom views based on LDAP search queries. Figure 5-73 shows the building of a query that generates the list of XIV Storage System accounts whose names start with XIV, and whose description is one of the following three: "Storage Administrator", "Read Only", or starts with "app".

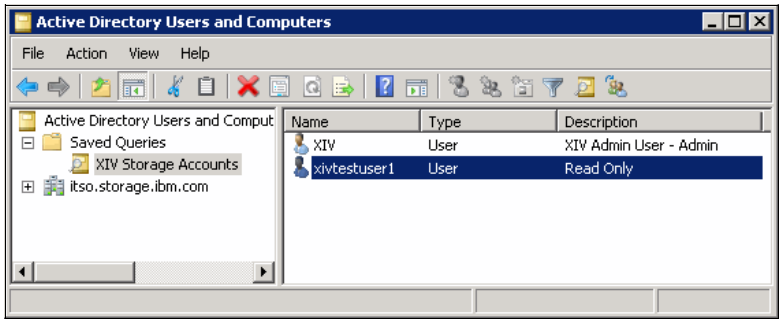


Figure 5-73 Active Directory query listing XIV accounts

To generate the XIV Storage System Accounts view, we used the LDAP query that is shown in Example 5-18.

Example 5-18 LDAP query for generating a list of XIV user accounts

```
(&(&(objectCategory=user)(cn=xiv*)(|(description=Read Only)(description=Storage
Administrator)(description=app*))))
```

To create this query, click **Saved Queries** → **New** → **Query** → **XIV Storage Accounts** and select the query name in this example. Click **Define Query**. Click **Find** → **Custom Search** → **Advanced** and paste the LDAP query from Example 5-18 into the Enter LDAP Query field.

When a new user account is created and its name and attributes satisfy the search criterion, this user account automatically appears in the XIV Storage Accounts view. Any LDAP XIV Storage Management GUI front end supporting the LDAP Version 3 protocol can be used for creating views and managing LDAP entries (XIV user accounts).

Table 5-4 provides a list of commands that *cannot* be used for user account management when LDAP authentication mode is active.

Table 5-4 XIV commands unavailable in LDAP authentication mode

XIV command
user_define
user_update
user_rename
user_group_add_user
user_group_remove_user

Authentication: When the XIV Storage System operates in LDAP authentication mode, user account creation, listing, modification, and removal functions are provided by the LDAP server.

The **user_list** command can still operate when LDAP authentication mode is active. However, this command shows only locally defined XIV user accounts and not LDAP accounts, as shown in Example 5-19.

Example 5-19 user_list command output in LDAP authentication mode

```
>> user_list show_users=all
Name                Category              Group    Active ...
xiv_development     xiv_development                yes
xiv_maintenance     xiv_maintenance                yes
admin               storageadmin                   yes
technician          technician                     yes
GPreadonly          readonly                       no
>> ldap_mode_get
Mode
Active
>>
```

As shown in Example 5-19 on page 270, the **Active** parameter is set to no for the GPReadonly user. The parameter specifies whether a user can log in to the current authentication mode. All predefined local XIV Storage System users can still log in when LDAP authentication mode is active.

Defining user groups with the GUI in LDAP authentication mode

User group information is stored locally on the XIV Storage System regardless of the authentication mode. The user group concept applies only to users assigned to an application_administrator role.

A user group can also be associated with one or multiple hosts or clusters.

To create user groups, add users (with application administrator role) to the group, and define host associations for the group, complete the following steps:

1. Be sure to log in as admin (or another user with storage administrator rights). From the Access menu, click **Users**, as shown in Figure 5-74. In our scenario, we create a user group called itso_app01_group. The user groups can be selected from the Access menu (padlock icon).

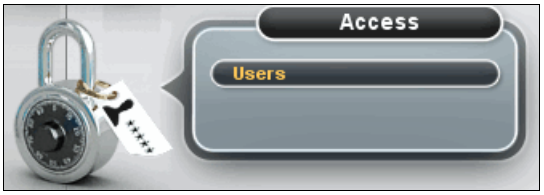


Figure 5-74 Select user groups

2. In the User Groups window, to add a user group, either click the **Add User Group** icon (shown in Figure 5-75) in the menu bar, or right-click in an empty area of the User Groups table and select **Add User Group** from the menu.

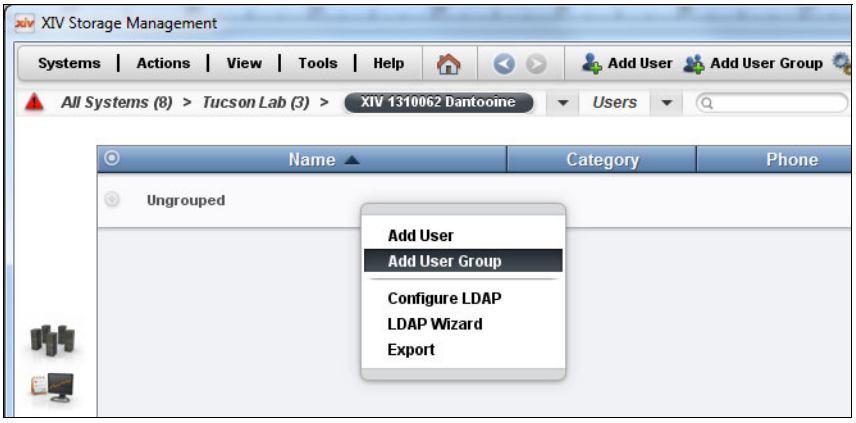


Figure 5-75 Add User Group selection

3. In the Add User Group window, enter a meaningful group name, specify a role for LDAP role mapping, as described in 5.7.4, “LDAP role mapping” on page 247, and click **Add** (Figure 5-76). To avoid potential conflicts with already registered user groups, the XIV Storage System verifies the uniqueness of the group name and the role. If a user group with the same name or the same role exists in the XIV Storage System repository, the attempt to create a user group fails and an error message is displayed.

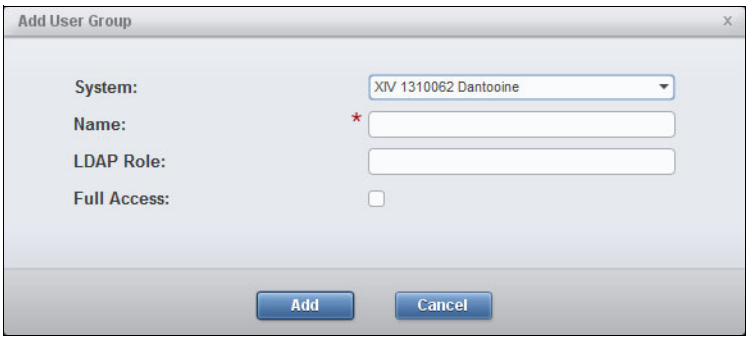


Figure 5-76 Enter new user group name and role for LDAP role mapping

The *Full Access* flag has the same significance as in native authentication mode. If a user group has the Full Access flag turned on, all members of that group have unrestricted access to all snapshots on the system.

At this stage, the user group `itso_app01_group` is still empty.

4. Next, add a host to the user group by right-clicking the name of the user group that you have created to open a menu and select **Update Access Control**, as shown in Figure 5-77.

Figure 5-77 Updating access control for a user group

The User Group Access Control window that is shown in Figure 5-78 opens. The window contains the names of all the hosts and clusters defined on the XIV Storage System.

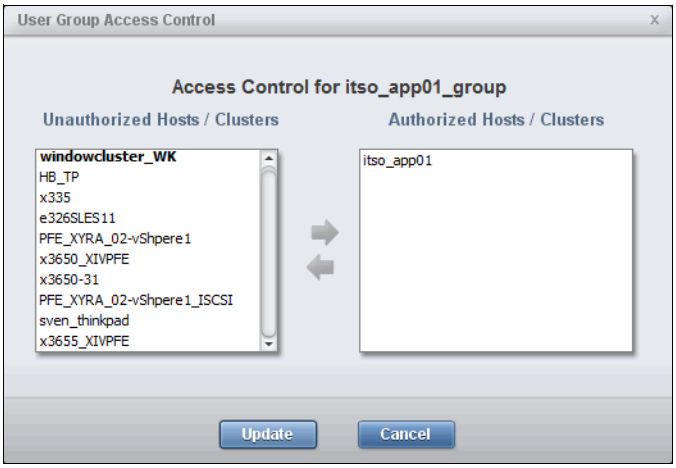


Figure 5-78 Access Control Definitions window

The left pane shows the list of unauthorized hosts and clusters for this particular user group, and the right pane shows the list of hosts that have already been associated with the user group.

5. Add or remove hosts from either list by selecting a host and clicking the appropriate arrow.
6. Finally, click **Update** to save the changes.

Unlike in native authentication mode, in LDAP authentication mode, user group membership cannot be defined using the XIV Storage Management GUI or XCLI. The group membership is determined at the time the LDAP authenticated user logs in to the XIV Storage System, based on the information stored in the LDAP directory. A description of the process of determining user group membership can be found in 5.7.4, “LDAP role mapping” on page 247.

Starting with XIV Storage System Software v10.2.2, it is possible to delete user groups when LDAP authentication is enabled.

5.9.3 Managing user groups by using the XCLI in LDAP authentication mode

Various commands and options are available to manage user groups, roles, and associated host resources through the XCLI.

To use the XIV Storage Management GUI to define user groups, complete the following steps:

1. Run **user_group_create**, as shown in Example 5-20, to create a user group called `itso_app02_group` with the corresponding LDAP role `itso_app01_admin`.

   Example 5-20 XCLI user_group_create in LDAP authentication mode

   ```
   >> user_group_create user_group=itso_app02_group
   ldap_role=CN=itso_app02_admin,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
   Command executed successfully
   ```

   Spaces: Avoid spaces in user group names. If spaces are required, the group name must be placed between single quotation marks, such as “name with spaces.”

   The `itso_app02_group` user group is empty and has no associated hosts or clusters. The next step is to associate a host or cluster with the group.

2. Associate the `itso_app02_group` user group to `host itso_app02`, as shown in Example 5-21.

   Example 5-21 XCLI access_define

   ```
   >> access_define user_group=itso_app02_group host=itso_app02
   Command executed successfully
   ```

5.9.4 Active Directory group membership and XIV role mapping

In Active Directory, a user can be a member of a single group or multiple groups. An LDAP group is a collection of users with common characteristics. *Group* is defined in the Active Directory container *Users*. A group is defined first as an empty container, and then existing users can be assigned as members of this group. A group is represented as a separate object in the LDAP DIT and gets a distinguished name (DN) assigned to it.

Groups defined in the Active Directory can be used for XIV Storage System role mapping. When a user becomes a member of a group in the Active Directory, it gets a new attribute assigned. The value of the new attribute points to the DN of the group. `MemberOf` is the name of that attribute. The `MemberOf` attribute value determines the Active Directory group membership.

To create a group in Active Directory, complete the following steps:

1. Start Active Directory Users and Computers by clicking **Start** → **Administrative Tools** → **Active Directory Users and Computers**.
2. Right-click the **Users** container and click **New** → **Group**.
3. Enter a group name and click **OK**.

The new Active Directory group creation window is shown in Figure 5-79.

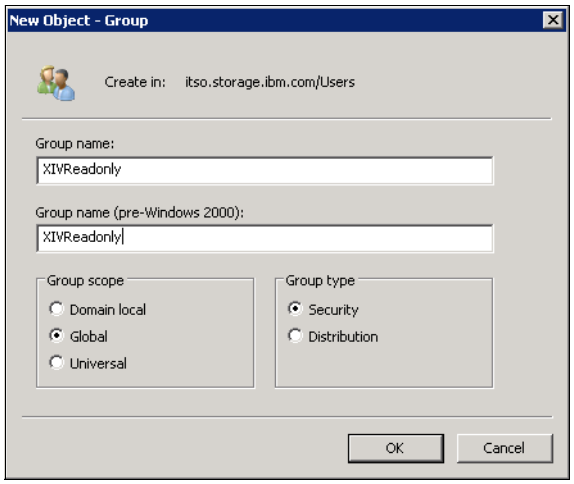


Figure 5-79 Creating Active Directory group

To assign an existing user to the new group, complete the following steps:

1. Start Active Directory Users and Computers by selecting **Start** → **Administrative Tools** → **Active Directory Users and Computers**.
2. Expand the **Users** container, right-click the user name that you want to make a member of the new group, and select **Add to a group**.
3. In the Select Groups window, click **Add** → **Advanced** → **Find Now**. From the presented list of existing user groups, click **XIVReadonly**, and then click **OK**.
4. You can now see a group selection window, as shown in Figure 5-80. Confirm your choice by clicking **OK**.

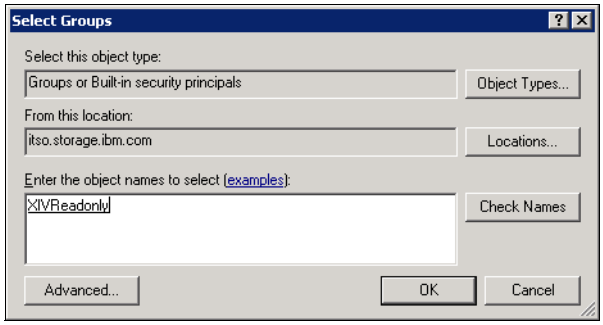


Figure 5-80 Active Directory group selection window

To illustrate the new `memberOf` attribute in the existing LDAP user object and the new LDAP object representing the “XIVReadOnly” group, we run **`ldapsearch`** queries against the Active Directory LDAP server, as shown in Example 5-22 on page 276.

Example 5-22 Active Directory group membership ldapsearch queries

```
# ldapsearch -LLL -H ldap://9.155.113.143:389 -D
cn=itso,cn=Users,dc=itso,dc=storage,dc=ibm,dc=com -w Passw0rd -b
cn=Users,dc=itso,dc=storage,dc=ibm,dc=com cn=itso_appuser1 memberOf
dn: CN=itso_appuser1,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
memberOf: CN=XIVReadonly,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com

# ldapsearch -LLL -H ldap://9.155.113.143:389 -D
cn=itso,cn=Users,dc=itso,dc=storage,dc=ibm,dc=com -w Passw0rd -b
cn=Users,dc=itso,dc=storage,dc=ibm,dc=com CN=XIVReadonly
dn: CN=XIVReadonly,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
objectClass: top
objectClass: group
cn: XIVReadonly
member: CN=itso_appuser1,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
member: CN=Almira,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
member: CN=Administrator,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
distinguishedName: CN=XIVReadonly,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
instanceType: 4
whenCreated: 20110831113851.0Z
whenChanged: 20111004132323.0Z
uSNCreated: 12856
uSNChanged: 41016
name: XIVReadonly
objectGUID:: g7EDY9DA1USZ6fTBfGbkCw==
objectSid:: AQUAAAAAAAUVAAAAZaIKWj+UHFDf+i81VQQAAA==
sAMAccountName: XIVReadonly
sAMAccountType: 268435456
groupType: -2147483646
objectCategory: CN=Group,CN=Schema,CN=Configuration,DC=itso,DC=storage,DC=ibm,
 DC=com
```

In the first **ldapsearch** query, we intentionally limited our search to the `memberOf` attribute (at the end of the **ldapsearch** command) so that the output is not obscured with unrelated attributes and values. The value of the `memberOf` attribute contains the DN of the group.

The second **ldapsearch** query illustrates the CN=XIVReadonly LDAP object content. Among other attributes, it contains the member attribute that points at the DN of the user defined as a member. The attribute member is a multivalued attribute; there can be more than one user assigned to the group as a member. `MemberOf` is also a multivalued attribute, and a user can be a member of multiple groups.

The XIV Storage System can now be configured to use the `memberOf` attribute for role mapping. In Example 5-23 on page 277, we map the Active Directory group XIVReadonly to the XIV read_only_role, XIVStorageadmin to the storage_admin_role, and XIV user group app01_group to Active Directory group XIVapp01_group. You must be logged on as admin.

Example 5-23 Configuring XIV to use Active Directory groups for role mapping

```
>> ldap_config_set xiv_group_attrib=memberOf
Command executed successfully.
>> ldap_config_set
read_only_role=CN=XIVReadonly,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
Command executed successfully .
>> ldap_config_set
storage_admin_role=CN=XIVAdmins,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
Command executed successfully.
>> ldap_config_get
Name                      Value
base_dn                   CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
xiv_group_attrib          memberOf
third_expiration_event    7
version                   3
user_id_attrib            objectSid
current_server            itso.storage.ibm.com
use_ssl                   no
session_cache_period      10
second_expiration_event   14
read_only_role            CN=XIVReadonly,DC=itso,DC=storage,DC=ibm,DC=com
storage_admin_role        CN=XIVAdmins,DC=itso,DC=storage,DC=ibm,DC=com
first_expiration_event    30
bind_time_limit           30

>> user_group_update user_group=itso_app01_group
ldap_role=cn=XIVAppAdmin1,DC=itso,DC=storage,DC=ibm,DC=com
Command executed successfully.
>> user_group_list user_group=itso_app01_group
Name               Access All   LDAP Role
Users
itso_app01_group   no           cn=XIVAppAdmin1,DC=itso,DC=storage,DC=ibm,DC=com
```

Alternatively, the same configuration steps can be accomplished by using the XIV Storage Management GUI. To change the LDAP configuration settings using the XIV Storage Management GUI, open the **Tools** menu at the top of the main XIV Storage Manager window, click **Configure** → **LDAP** → **Role Mapping**, and change the configuration parameter settings.

Now, by assigning Active Directory group membership, you can grant access to the XIV Storage System.

A user in Active Directory can be a member of multiple groups. If this user is a member of more than one group with corresponding role mapping, XIV fails authentication for this user because the role cannot be uniquely identified. In Example 5-24 on page 278, the xivtestuser1 user can be mapped to Storage Admin and Read Only roles, which is why the authentication failure followed by the `USER_HAS_MORE_THAN_ONE_RECOGNIZED_ROLE` error message occurs.

Example 5-24 LDAP user mapped to multiple roles authentication failure

```
$ xcli -c "XIV 6000050" -u xivtestuser1 -p Passw0rd ldap_user_test
Error:   USER_HAS_MORE_THAN_ONE_RECOGNIZED_ROLE
Details: User xivtestuser has more than one recognized LDAP role.

$ ldapsearch -LLL -H ldap://9.155.113.143:389 -D
cn=itso,cn=Users,dc=itso,dc=storage,dc=ibm,dc=com -w Passw0rd -b
cn=Users,dc=itso,dc=storage,dc=ibm,dc=com cn=xivtestuser1 memberOf

dn: CN=xivtestuser1,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
memberOf: CN=XIVReadonly,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
memberOf: CN=XIVAdmins,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
```

An LDAP user can be a member of multiple Active Directory groups and successfully authenticate to an XIV Storage System if only one of those groups is mapped to an XIV Storage System role. As illustrated in Example 5-25, the xivtestuser2 user is a member of two Active Directory groups: XIVAdmins and NonXIVgroup. Only XIVAdmins is mapped to an XIV role.

Example 5-25 LDAP user mapped to a single roles authentication success

```
$ xcli -c "XIV 6000050" -u xivtestuser2 -p Passw0rd ldap_user_test
Command executed successfully.

$ ldapsearch -LLL -H ldap://9.155.113.143:389 -D
cn=itso,cn=Users,dc=itso,dc=storage,dc=ibm,dc=com -w Passw0rd -b
cn=Users,dc=itso,dc=storage,dc=ibm,dc=com cn=xivtestuser2 memberOf

dn: CN=xivtestuser2,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
memberOf: CN=NonXIVgroup,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
memberOf: CN=XIVAdmins,CN=Users,DC=itso,DC=storage,DC=ibm,DC=com
```

After all Active Directory groups are created and mapped to corresponding XIV Storage System roles, the complexity of managing LDAP user accounts is reduced because the role mapping can now be done through Active Directory group membership management. The easy to use point-and-click interface leaves less room for error when it comes to assigning group membership, as opposed to entering text into the description field.

5.9.5 OpenLDAP Directory and XIV Storage System role mapping

OpenLDAP Directory group membership can be used for XIV Storage System role mapping as described in 5.9.4, “Active Directory group membership and XIV role mapping” on page 274 by using the `memberOf` attribute to explicitly identify what group memberships that a user has.

The role mapping mechanism is not restricted to using group membership to identify association to a suitable XIV Storage System role. It can use a text-based attribute directly identifying the XIV Storage System role the user has been associated with.

Configuring the XIV Storage System for OpenLDAP Directory

Configuration for the OpenLDAP Directory is essentially the same as other directories. The only difference is that we use the *description* attribute with OpenLDAP. If the description attribute is already in use, then you can substitute any suitable unused text attribute.

The schema in our example has been built with the same DN base, but OpenLDAP has some subtle differences. For example, individual user objects are defined as UID type objects as opposed to using CN type with Active Directory. Also, the “Users” group is a OU type object, as shown in Example 5-26.

Example 5-26 Listing objects in an OpenLDAP schema

```
# ldapsearch -LLL -b ou=Users,dc=itso,dc=storage,dc=ibm,dc=com -D
cn=Manager,dc=itso,dc=storage,dc=ibm,dc=com -w PASSWOrd OU=Users
dn: ou=Users,dc=itso,dc=storage,dc=ibm,dc=com
objectClass: top
objectClass: organizationalUnit
ou: Users

# ldapsearch -LLL -b ou=Users,dc=itso,dc=storage,dc=ibm,dc=com -D
cn=Manager,dc=itso,dc=storage,dc=ibm,dc=com -w PASSWOrd dn
dn: ou=Users,dc=itso,dc=storage,dc=ibm,dc=com
dn: uid=itso,ou=Users,dc=itso,dc=storage,dc=ibm,dc=com
dn: uid=zejn,ou=Users,dc=itso,dc=storage,dc=ibm,dc=com
dn: uid=almira,ou=Users,dc=itso,dc=storage,dc=ibm,dc=com
dn: cn=XIVAdmins,ou=Users,dc=itso,dc=storage,dc=ibm,dc=com
dn: cn=XIVReadonly,ou=Users,dc=itso,dc=storage,dc=ibm,dc=com
dn: uid=phb,ou=Users,dc=itso,dc=storage,dc=ibm,dc=com
```

The XIV administrator must specify that the LDAP server type is “Open Ldap”, using the **ldap_config_set** XCLI command as follows:

```
ldap_config_set server_type=”Open Ldap”
```

This command also automatically sets both the **user_id_attrib** and **user_name_attrib** values to “uid”.

Then the `xiv_group_attrib` configuration value is changed to a description as follows:

```
ldap_config_set xiv_group_attrib=description
```

The role names must be set to reflect the storageadmin and readonly roles. Because there are no group mappings, they do not need to have the full DN path as when using group membership. Instead, we need to use only “XIVAdmins” and “XIVReadonly” values for each role:

- **ldap_config_set storage_admin_role=”XIVAdmins”**
- **ldap_config_set read_only_role=”XIVReadonly**

The LDAP service user and password must be defined. This user is a user that needs no specific permissions, within the directory, and is only used to view the directory schema. Run the following command:

```
ldap_config_set xiv_user=“UID=XIV,OU=Users,DC=itso,DC=storage,DC=ibm,DC=com”
xiv_password=PasswOrd
```

Then, the LDAP server instance must be defined using the **ldap_add_server** command as follows:

```
ldap_add_server fqdn=itso.storage.ibm.com port=389 address=9.155.113.137
base_dn="OU=Users,DC=itso,DC=storage,DC=ibm,DC=com"
```

In Example 5-27 on page 280, we apply these changes to allow the OpenLDAP connection using the XCLI. You must be logged in to the XCLI as the admin user.

Example 5-27 Configuring XIV Storage System to use OpenLDAP Directory groups for role mapping

```
>>ldap_config_set server_type="Open Ldap"
Command executed successfully.

>> ldap_config_set xiv_group_attrib=description
Command executed successfully.

>> ldap_config_set read_only_role=XIVReadonly
Command executed successfully.

>> ldap_config_set storage_admin_role=XIVAdmins
Command executed successfully.

>> ldap_config_set xiv_user="UID=XIV,OU=Users,DC=itso,DC=storage,DC=ibm,DC=com"
xiv_password=PasswOrd
Command executed successfully.

>>ldap_add_server fqdn=itso.storage.ibm.com port=389 address=9.155.113.137
base_dn="OU=Users,DC=itso,DC=storage,DC=ibm,DC=com"
Command executed successfully.

>>ldap_config_get
Name                          Value
current_server
version                       3
xiv_group_attrib              description
storage_admin_role            XIVAdmins
read_only_role                XIVReadonly
session_cache_period          20
bind_time_limit               20
user_id_attrib                uid
first_expiration_event        30
second_expiration_event       14
third_expiration_event        7
use_ssl                       no
xiv_user                      UID=XIV,OU=Users,DC=itso,DC=storage,DC=ibm,DC=com
server_type                   Open Ldap
user_name_attrib              uid
group_search_depth            0
group_search_max_queries      39
group_search_stop_when_found  yes

>>ldap_list_servers
FQDN                  Address         Base DN ...
itso.storage.ibm.com  9.155.113.137   OU=Users,DC=itso,DC=storage,DC=ibm,DC=com
```

Alternatively, the same configuration steps can be accomplished by using the XIV Storage Management GUI. To change the LDAP configuration settings in the XIV Storage Management GUI, open the **Tools** menu at the top of the main XIV Storage Manager window, click **Configure** → **LDAP** → **Role Mapping**, and change the configuration parameter settings.

Associating roles in OpenLDAP

Because we are just using a text string to identify a user in an XIV Storage System group role, the string must be added to the description attribute.

In Example 5-28, we add the “XIVAdmins” role association to the itso user. Because we are using a command-line utility, a text file is used to feed in the relevant commands into the **ldapmodify** command.

Example 5-28 Adding role association to an OpenLDAP user

```
# cat user_mod_xivadmins.itso
dn: uid=itso,ou=Users,dc=itso,dc=storage,dc=ibm,dc=com
changetype: modify
replace: description
description: XIVAdmins

# ldapmodify -x -r -f  user_mod_xivadmins.itso -D
cn=Manager,dc=itso,dc=storage,dc=ibm,dc=com -w PASSWOrd -v
ldap_initialize( <DEFAULT> )
replace description:
        XIVAdmins
modifying entry "uid=itso,ou=Users,dc=itso,dc=storage,dc=ibm,dc=com"
modify complete

# ldapsearch -LLL -b ou=Users,dc=itso,dc=storage,dc=ibm,dc=com -D
cn=Manager,dc=itso,dc=storage,dc=ibm,dc=com -w PASSWOrd uid=itso description
dn: uid=itso,ou=Users,dc=itso,dc=storage,dc=ibm,dc=com
description: XIVAdmins
```

We can assign the “XIVReadonly” value to the description attribute (or any other defined application admin LDAP role in the XIV).

We can now test (with XCLI) if that user can authenticate with the LDAP server and has the correct XIV role mapping, as shown in Example 5-29.

Example 5-29 Testing if an LDAP user has the correct XIV role mapping permission

```
>>ldap_test user=itso password=PasswOrd fqdn=itso.storage.ibm.com
Command executed successfully.
```

If a user does not have a value (or an incorrect value) in the description attribute, then an appropriate error is displayed, as shown in Example 5-30.

Example 5-30 Testing an LDAP user without correct XIV role mapping permission

```
>>ldap_test user=PHB password=nO1dea fqdn=itso.storage.ibm.com
Error:   LOGIN_FAILURE_USER_MISSING_GROUP_ATTRIBUTE
Details: User PHB is missing the group attribute 'description'.
>>ldap_test user=Wally password=KOffee fqdn=itso.storage.ibm.com
Error:   LOGIN_FAILURE_USER_HAS_NO_RECOGNIZED_ROLE
Details: User Wally has no recognized LDAP role.
```

Activating LDAP authentication

After you can confirm LDAP users are tested successfully with the correct LDAP permissions, you can enable LDAP authentication by changing the XIV Storage System LDAP mode using the **ldap_set_mode** XCLI command, as shown in Example 5-31 on page 282.

Example 5-31 Enabling LDAP in XIV Storage System

```
>>ldap_mode_set mode=Active

Warning:   ARE_YOU_SURE_YOU_WANT_TO_ENABLE_LDAP_AUTHENTICATION y/n:
Command executed successfully.
```

To disable access, repeat the same command with the **mode=Inactive** option.

5.9.6 Managing multiple systems in LDAP authentication mode

The task of managing multiple XIV systems can be simplified by using LDAP authentication mode. Because user credentials are stored centrally in the LDAP directory, it is no longer necessary to synchronize user credentials among multiple XIV systems. After a user account is registered in LDAP, multiple XIV systems can use credentials stored in LDAP directory for authentication.

Because the user's password is stored in the LDAP directory, all connected XIV systems authenticate the user with this password. If the password is changed, all XIV systems automatically accept the new password. This mode of operation is often referred to as *single sign-on* (SSO) (Figure 5-81). This mode allows for quick transitions between systems in the XIV Storage Management GUI because the password is entered only once. This approach is especially useful in Remote Mirror configurations, where the storage administrator is required to frequently switch from source to target system.

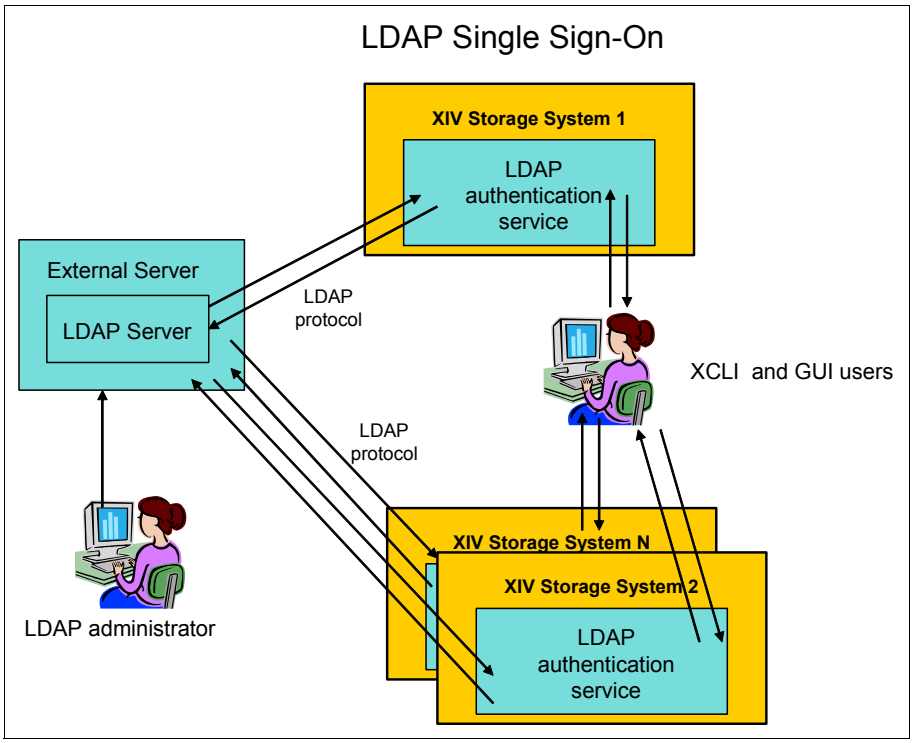


Figure 5-81 LDAP single sign-on

Important: To allow single sign-on in LDAP authentication mode, all XIV systems must be configured to use the same set of LDAP configuration parameters for role mapping. If role mapping is set up differently on any two XIV systems, it is possible that a user can log in to one but not the other system.

5.10 Securing LDAP communication with Secure Sockets Layer

In any authentication scenario, information is exchanged between the LDAP server and XIV Storage System where access is being sought. Secure Sockets Layer (SSL) can be used to implement secure communications between the LDAP client and server. LDAP over SSL (LDAPS), the secure version of the LDAP protocol, allows secure communications between the XIV Storage System and LDAP server with encrypted SSL connections. This protocol allows a setup where user passwords never appear on the wire in clear text.

SSL provides methods for establishing identity using X.509 certificates and ensuring message privacy and integrity using encryption. To create an SSL connection, the LDAP server must have a digital certificate signed by a trusted certificate authority (CA). Companies have the choice of using a trusted CA from another vendor or creating their own certificate authority. In this scenario, the itso.storage.ibm.com CA is used for demonstration purposes.

5.10.1 Configuring XIV to use LDAP over Secure Sockets Layer

To be operational, SSL must be configured on both the client (XIV Storage System) and the server (LDAP server). Server configuration includes generating a certificate request, obtaining a server certificate from a CA, and installing the server and CA certificates. The `cacert.pem` file is ready to be uploaded to the XIV Storage System.

When a new LDAP server is added to the XIV Storage System configuration, a security certificate can be entered in the optional certificate field. If the LDAP server was originally added without a certificate, you must remove that definition first and add a definition with the certificate.

LDAP server: When defining the LDAP server with a security certificate in XIV Storage System, the fully qualified name of the LDAP server must match the “issued to name” in the client’s certificate.

For registering the LDAP server with a security certificate, it might be easier to use the XIV Storage Management GUI, because it has a file upload capability (see Figure 5-82 on page 284). XCLI can also be used, but in this case you need to cut and paste a long string containing the certificate into the XCLI session. To define the LDAP server in the XIV Storage Management GUI, from the menu bar, click **Systems** → **Settings** → **LDAP**, click the **LDAP Servers** tab, and then click the green plus sign (**+**) on the right panel.

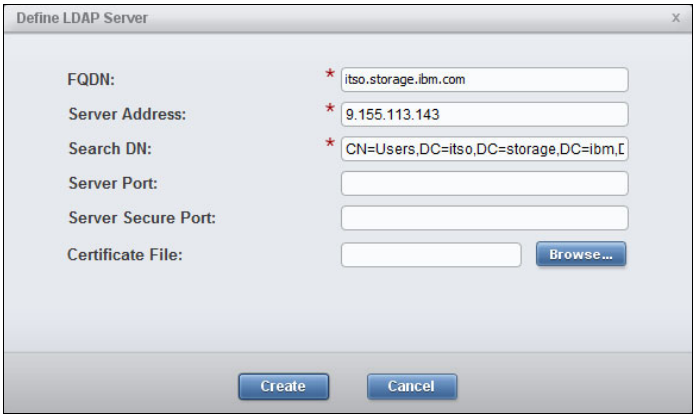


Figure 5-82 Defining Active Directory LDAP server with an SSL certificate

In Figure 5-82, the server type selected must correspond to your specific LDAP directory, either Microsoft Active Directory, as shown, or an Oracle directory.

To view the expiration date of the installed certificate in the XIV Storage Management GUI, open the **Tools** drop-down menu at the top of the main XIV Storage Manager GUI window. Click **Settings** → **LDAP**, click the **Servers** tab, right-click the name of the LDAP server, and click **Properties**, as shown in Figure 5-83.

Figure 5-83 Viewing the Active Directory server certificate expiration date

By default, LDAP authentication on an XIV Storage System is configured to use non-SSL communication. To enable the use of SSL in the XIV Storage Management GUI, open the **Tools** drop-down menu at the top of main XIV Storage Manager window. Click **Settings** → **LDAP**, click the **Secure LDAP** tab, and change the Use SSL setting from No to **Yes**, as shown in Figure 5-84 on page 285.

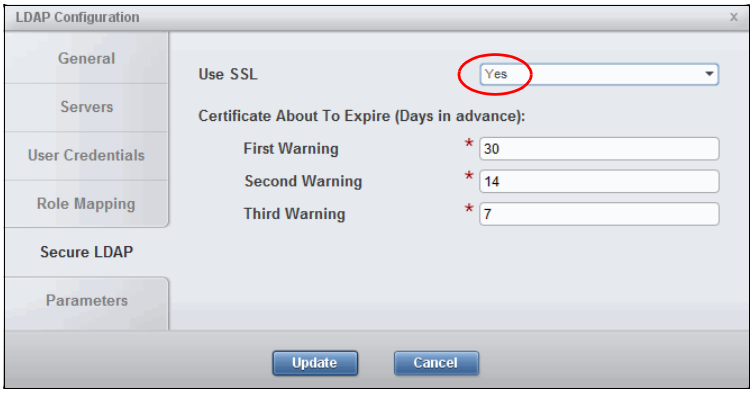


Figure 5-84 Enabling SSL for Active Directory LDAP communication

5.10.2 Maintaining the SSL certificate

A new SSL certificate must be installed before the existing one expires. If you let your SSL certificate expire, XIV Storage System LDAP authentication is no longer possible until you either disable SSL or install the new certificate on both the LDAP server and the XIV Storage Server. Before the SSL certificate expires, the XIV Storage System issues three notification events. The first “LDAP SSL Certificate About to Expire” event is shown in Figure 5-85.

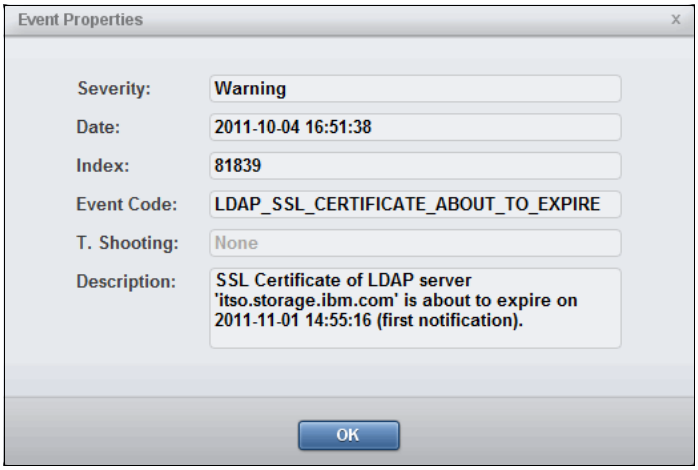


Figure 5-85 First notification of SSL certificate of LDAP server expiration

6

Performance

This chapter describes how the IBM XIV Storage System software and hardware work together to provide high performance characteristics. These characteristics, which are inherent in the XIV Storage System design, deliver optimized and consistent performance. There is little that clients need to do within an XIV Storage System that can contribute to performance gains beyond what the XIV Storage System automatically provides, and that includes the optional flash caching (solid-state drive (SSD) caching) feature. However, there are several practices *from a host perspective* that can have a positive impact on XIV Storage System performance.

In addition, this chapter describes XIV Storage System performance monitoring techniques using the IBM XIV Storage Management graphical user interface (GUI) and XIV Top. Performance problem-solving techniques are also described. We provide guidance about how to interpret XIV Storage System performance information to determine objective levels of XIV Storage System performance characteristics such as response times, I/O rates, and throughput.

Performance is one of the primary strengths of the XIV Storage System. This chapter covers the following topics as they pertain to XIV performance:

- 6.1, “XIV Storage System Software and hardware architecture” on page 288
- 6.2, “Practices for optimum performance” on page 294
- 6.3, “Performance monitoring” on page 304
- 6.4, “Performance evaluation” on page 320

A popular topic that this chapter does not cover is the actual performance feeds and speeds of the XIV Storage System. That information is available in the IBM System Storage XIV Performance white paper:

http://public.dhe.ibm.com/common/ssi/ecm/en/tsw03123wwen/TSW03123WWEN.PDF

More white papers about XIV performance with application-specific workloads are available in the White Papers section of the IBM Techdocs Library:

http://www.ibm.com/support/techdocs/atsmastr.nsf/Web/WhitePapers

© Copyright IBM Corp. 2014. All rights reserved.

6.1 XIV Storage System Software and hardware architecture

There are several aspects of the XIV Storage System hardware and software architecture that contribute to the ability of an XIV Storage System to provide excellent performance to real-world production applications. In this section, we describe the following topics:

- Workload distribution
- Grid architecture
- Caching and extended flash caching
- Data redistribution
- Snapshots

6.1.1 Workload distribution and load balancing

One of the main reasons that the XIV Storage System achieves exceptional performance is because the architecture imposes an elegant way to equally and in parallel use every component within the XIV Storage System grid. Equal use of every component maximizes performance in three ways:

- XIV Storage System engages the performance capabilities of *all drives all the time*.
- XIV Storage System engages the performance capabilities of *every module* in the grid.
- XIV Storage System *eliminates disk hotspots*.

The XIV Storage System stripes and mirrors data in 1 MB partitions across all of the disks in the system. It disperses the 1 MB partitions in a pseudo-random distribution. For more details about the architecture of the system, see Chapter 2, “IBM XIV Storage System logical architecture and concepts” on page 17.

For disk systems, one of the most important contributors to performance is the number of disk drives that can simultaneously work together to handle an application’s I/O requirements. Traditionally, this situation has been limited by the number of disk drives that can be connected to Redundant Array of Independent Disks (RAID) controller pairs. As we know, XIV Storage System does not follow this traditional approach. Instead, XIV Storage System evenly distributes every logical unit number (LUN) assigned to an application across every disk drive in every module of the XIV Storage System. This situation is true for every XIV Storage System LUN size and every XIV Storage System module configuration. This situation means that, on average, an application server is using every XIV Storage System disk drive equally. To emphasize this point, see Figure 6-1, which shows a LUN striped across a traditional RAID 5 array that consists of eight disk drives.

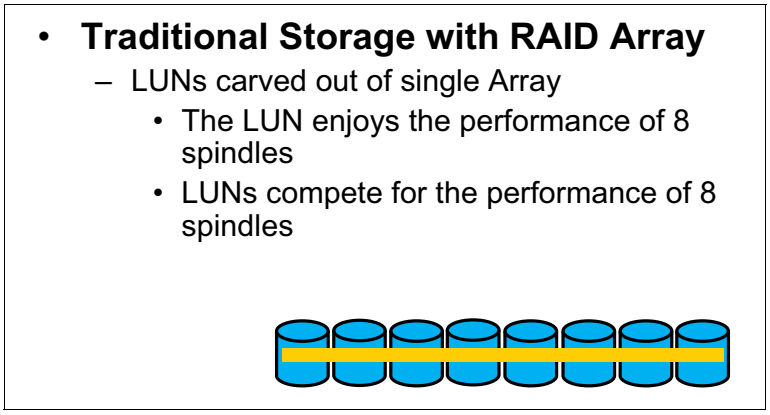

• Traditional Storage with RAID Array
– LUNs carved out of single Array
• The LUN enjoys the performance of 8 spindles
• LUNs compete for the performance of 8 spindles

Figure 6-1 Eight-drive RAID array

Now, in contrast, see Figure 6-2, which shows a LUN virtualized into 1 MB partitions and distributed evenly across 180 drives in a 15-module XIV Storage System configuration.

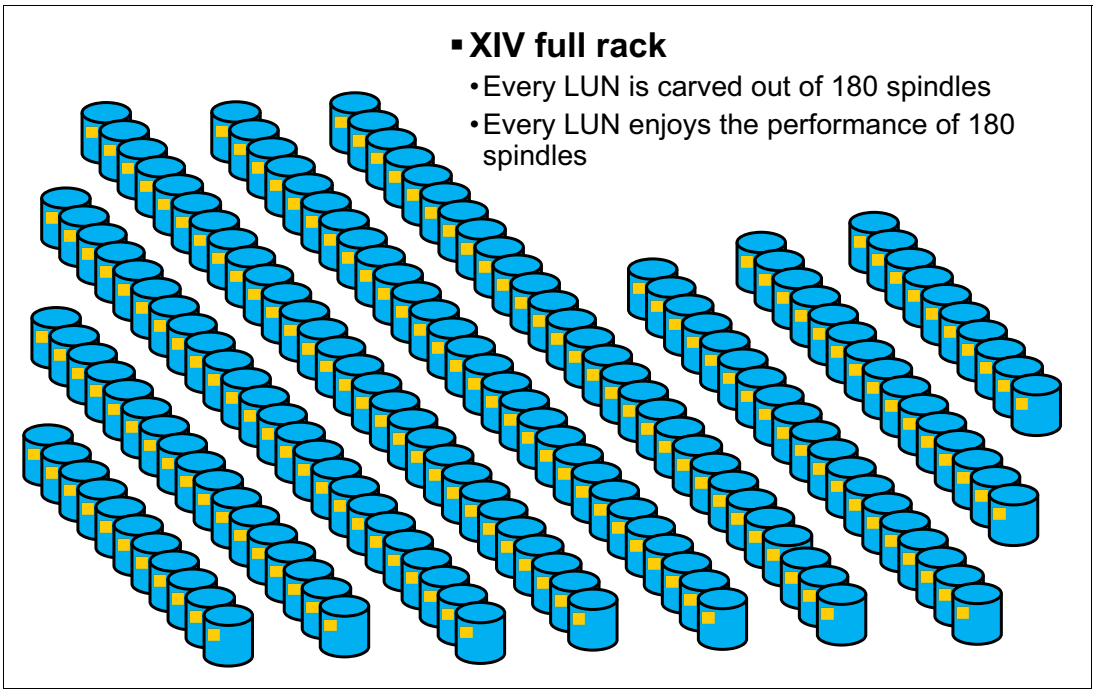


Figure 6-2 XIV Storage System full rack with 180 drives

In a traditional storage system, it is common for an application's performance requirements to be greater than the performance capabilities of a single RAID array. To engage multiple traditional RAID arrays into providing the performance requirements of an application, one or more of the following techniques must be used:

- Break up the application into multiple LUNs and distribute them among several RAID arrays
- Use operating system logical volume striping techniques to distribute the operating system across the performance capabilities of multiple RAID arrays
- Storage subsystem striping techniques that create LUNs that are striped across multiple RAID arrays

It is common for these traditional storage system performance techniques to be combined to meet the performance requirements of an application. These techniques are a necessity with traditional storage. They can become challenging if there is ever the need to change these configurations because more storage capacity or more performance capacity is required. These activities require specialized skills and much work.

With XIV Storage System, there is no need to resort to these multiple techniques. Each XIV Storage System LUN has only one disk performance limitation and that is the full performance potential of the entire XIV Storage System, which includes all disks and modules in the XIV Storage System configuration. The consequence of this massive parallel approach and the way data is effectively laid out within the XIV Storage System grid is that LUNs can be easily resized without physical disk layout concerns. It is unlikely that an XIV Storage System LUN will *ever* be reconfigured for performance reasons.

What if you add more than one application to a host server and consider multiple host systems in a typical client computing environment? For traditional storage systems where RAID arrays, or groups of arrays, are dedicated to an application, the considerations for providing the adequate number of disk drives to handle the application performance requirements increase linearly with the number of applications. These groups of arrays become islands within the traditional storage system. Realistically, several of these islands are sized correctly for the application's performance requirements, where others are oversized to handle application performance growth projections. The islands that are undersized are a performance bottleneck problem that must be addressed.

For the XIV Storage System, there are no special considerations for multiple applications on multiple application servers. If there is adequate spare performance capacity in the XIV Storage System to handle the performance requirements of the new application, all that is required is to create new LUNs and assign them to the host. However, if the XIV system is to be used to the maximum of the performance that it can deliver, you can use the quality of service (QoS) feature to ensure that your business-critical applications are served in priority with adequate performance, from an input/output operations per second (IOPS) and bandwidth standpoints (for details, see 4.7, "QoS feature" on page 170).

Because of the grid architecture, the XIV Storage System actually works more efficiently as multiple application servers are attached. The XIV evenly distributes the production workload over the performance resources of the multiple XIV Storage System interface modules and, therefore, the entire grid architecture.

6.1.2 Grid architecture

Each XIV Storage System module contains disk drives that contribute to the data storage capacity and performance capacity of the XIV Storage System configuration. If there is enough available performance capacity within the XIV to handle the new application's performance requirements, the only configuration activity required is to create new LUNs. This situation assumes that there is also enough available data storage capacity. But what are the considerations if the XIV does not have additional performance capacity to take on more applications? Later in "Performance analysis" on page 321, we describe how you determine whether the XIV Storage System has adequate performance capacity to handle more workload.

Each XIV Storage System module is carefully designed and has the appropriate processing mechanisms to handle the I/O workload of the 12 disk drives. There are several aspects to this module I/O architecture.

A cache communicates to each of two serial-attached SCSI (SAS) disk controllers over multiple lanes of the module PCI Express (PCIe) bus. The SAS controllers themselves have the appropriate number of processing capabilities to handle the I/O requirements of the disk drives that they control. These controllers, in turn, communicate to the disk drives over multiple lanes of the PCIe bus. All of this communication is handled by the correctly sized processor. The cache is also correctly sized so that it can handle the total amount of I/O that each module can sustain at peak performance. Furthermore, the XIV Gen3 with the XIV System software Version 11.1.0 or later, also offers the option of an extended SSD cache. The SSD cache can boost performance to up to three times for typical online transaction processing (OLTP) workloads, with no additional configuration or tuning needed.

Note: SSD performance has been further enhanced in Version 11.2 of the XIV Storage System Software through optimization of SSD data integrity checking.

This module architecture is explored from a different perspective. Would it be possible for the XIV Storage System module to have one of the most powerful multi-core processors available today? Would it be possible for each module to have more cache or faster disk controllers or faster bus speeds or faster disk drives? The answer to all these questions is *yes*. But what is the point in having the latest, fastest components available if the total system can drive these components to only a fraction of their performance capabilities? The XIV Storage System module is carefully designed to use the performance capacity of individual components in a balanced way as the total system approaches its peak performance capabilities.

As an example, consider the new 6-core CPUs introduced in XIV Gen3 Model 214. There was a need to upgrade the CPU with more processing capacity to support the 10 Gb IP network Small Computer System Interface (iSCSI) available with XIV v11.2 software.

When a partial XIV Storage System configuration needs more performance capacity, it is a simple matter of adding more modules to the XIV Storage System configuration. When modules are added, the XIV Storage System Software automatically redistributes the existing data. The XIV system has a maximum of 15 modules.

This data redistribution has many interesting aspects with positive performance implications:

- Only actual data is redistributed. XIV Storage System is data-aware, which means that the work of redistribution is necessary only on partitions that contain actual data. Partitions that do not contain data do not take up any XIV Storage System performance resources during redistribution.
- All disk drives in the original configuration contribute equally to the redistribution activity. This situation results in the fastest redistribution times possible and minimizes stress on the drives.

One of the most impressive and unique aspects of the XIV Storage System grid architecture is that, as more modules are added to the system, performance basically scales up linearly. Again, this situation is because each module contains a balanced amount of additional processors, cache, disk performance, and bandwidth capacity.

6.1.3 Caching mechanisms

From a total system perspective, the XIV Storage System has ample cache. The XIV Storage System Gen3 Model 214 system with 4 TB drives, for example, was announced with 48 GB of cache per module for a fully configured system cache size of 720 GB. The XIV Gen3 with the XIV System software Version 11.1.0 or later also offers the option of an extended SSD cache. This extension provides an additional 6 TB of read cache (on a 15-module system). Since XIV System software Version 11.4.0, the read cache extension with flash drives can be up to 12 TB (on a 15-module system). The flash (SSD) cache can boost performance up to three times for typical OLTP workloads, with no additional configuration or tuning needed. The net result of this much cache is that the XIV Storage System can be aggressive about prestaging data into cache so that host system reads can be satisfied much faster than if the data must be retrieved from disk.

This section describes how the XIV Storage System global caching mechanisms work. For details about the flash (SSD) caching, see the Redpaper publication, *Solid-State Drive Caching in the IBM XIV Storage System*, REDP-4842.

The XIV Storage System distributes cache into each module. This distributed cache enables each module to concurrently service host I/Os and cache-to-disk access. Each module's cache just handles the movement of data into and out of that particular module. The processing work that one module does to move data into and out of cache has absolutely no adverse effect on any other module. Because of the XIV Storage System grid architecture,

this cache-related I/O workload is distributed evenly between all the XIV Storage System modules. This situation is different from other enterprise storage systems, which use a central memory cache that implements memory locking algorithms that generate access contention.

To improve memory management, each data module uses a PCIe bus between the cache and the disk controllers to provide a sizable connection between the disk and the cache. This design aspect allows the bus to transfer large amounts of data quickly between the disks.

Having a large bus bandwidth allows the XIV Storage System to use small cache pages. This large bus bandwidth between the disk and the cache also allows the system to perform many small requests in parallel, again improving the performance.

A least recently used (LRU) algorithm is the basis for the cache management algorithm. This feature allows the system to generate a high hit ratio for frequently used data. The efficiency of the cache use for small transfers is high when the host is accessing the same data set.

The cache algorithm starts with a single 4 KB page and gradually increases the number of pages prefetched until an entire partition, 1 MB, is read into cache. If the access results in a cache hit, the algorithm doubles the amount of data prefetched into the system.

The prefetching algorithm continues to double the prefetch size until a cache miss occurs or the prefetch size maximum of 1 MB is obtained. Because the modules are managed independently, if a prefetch crosses a module boundary, the logically adjacent module (for that volume) is notified to begin pre-staging the data into its local cache.

6.1.4 Data redistribution effects on host systems

This section describes the effects that the application server experiences during XIV Storage System data rebuild and redistribution. For our purposes here, we refer to these rebuild and redistribution actions as *data redistribution* because we are focused on the effects that these redistribution actions have on the storage I/O performance seen by the application server. For a fuller description of data rebuild and redistribution, read 2.10.2, “Preserving data redundancy: Rebuilding and redistributing” on page 51.

When, for instance, an XIV Storage System disk fails, the storage system takes that drive offline, identifies the actual data partitions that the drive contained, and locates the copy of those data partitions within the rest of the XIV drives. The XIV Storage System then makes a copy of only that actual data and moves that data evenly across each of the remaining disk drives. The performance impact during this data redistribution is negligible to the I/O characteristics that an XIV can provide to application servers.
A redistribution will *not* take place for a flash drive failure because the optional flash drives are not involved in data retention or redundancy. They are merely used as read cache.

Consider a fully populated XIV Storage System that experiences this drive failure. In this example, the XIV reads data from 168 drives. Each drive contains an equal portion of the mirror data copies of the data from the failed drive. The XIV reads data from 168 drives because there is no data from the failed drive in the same XIV module that contains the failed drive. The amount of data from the failed drive on each of the remaining 168 drives represents a small percentage of the total data on each drive. The XIV makes a copy of this data and redistributes it on reserved space on the remaining 179 drives. So this operation reads a small amount of data from 168 drives and, in turn, writes the same small amount of data to 179 drives. This operation represents many spindles doing little work to recover from a drive failure.

Consider the following observations about XIV Storage System data redistribution:

- Additional drive use is small, which means that applications hardly notices any I/O performance differences.
- The speed at which an XIV Storage System can redistribute data is fast because of all the disk spindles involved, resulting in relatively short rebuild and redistribution times as compared to RAID rebuild times.
- There is no concern for data access patterns because all XIV Storage System drives equally share the production I/O workload and redistribution workload.
- The redistribution process also has the advantage of massive parallel computing because each module works in parallel to manage only 12 drives and to generate data copies as quickly as possible.
- An XIV Storage System is data-aware, which means that XIV protects or redistributes *only* actual data that has been written. XIV Storage System does not bother protecting empty or free capacity that exists within a failed drive or module. This situation is different from how RAID-based arrays handle rebuilds and is one of the reasons why XIV Storage System rebuilds are fast.

It is easy to extend this to a failed XIV module because the same redistribution scenario takes place. The only difference is that now the data redistribution involves 12 drives rather than just one. All the concepts and considerations and advantages still apply. The rebuild and redistribution take longer because more data is involved.

The IOPS and megabytes per second (MBps) experience basically no change in the performance supplied to the applications.

Compare XIV Storage System redistribution to a RAID 5 array failure and rebuild:

- The entire data storage capacity must be rebuilt using parity recalculation, not just actual user data.
- The rebuild process imposes a significant stress on the drives, which can lead to additional drive failures.
- The rebuild has a significant impact on the performance capabilities of the array.
- Rebuild times are typically much longer when compared to XIV Storage System redistribution.

XIV Storage System redistribution is the same mechanism that redistributes data when a module is added to an XIV Storage System, with the same considerations. And it is worth noting that, although it has nothing to do with performance, XIV Storage System data redistribution is carried out automatically. No user intervention is necessary or even possible.

6.1.5 Snapshots

The performance penalty during an XIV Storage System snapshot is nearly invisible to the I/O characteristics of production application workload. Snapshots complete near instantly within the XIV Storage System. When a snapshot is issued, no data is copied. The snapshot creates system pointers to the original data. As the host writes modified data in the master volume, the XIV Storage System redirects the write data to a new partition. Only the data that was modified by the host is copied to the new partition. This prevents moving the data multiple times and simplifies the internal management of the data. See *IBM XIV Storage System Business Continuity Functions*, SG24-7759 for details about how the snapshot function is implemented.

6.2 Practices for optimum performance

There are several practices that can have improve XIV Storage System performance.

Important: XIV architecture is designed to support real-world client production environments.

Real-world production environments involve multiple application servers making multiple simultaneous I/O demands on storage systems.

When clients decide to purchase the XIV Storage System, they have the reasonable expectation that they will migrate existing applications or install new applications to the XIV Storage System and experience great performance.

In this section, we describe the following topics:

- Sizing
- Number of LUNs
- Multipathing considerations
- Host considerations:
  - Application threads
  - Host bus adapter (HBA) queue depth
  - Logical volume manager striping
  - Operating system tunables
- Quality of service (QoS)

6.2.1 Sizing

By far the most important aspect of XIV Storage System performance is to properly size the system based on performance requirements. Each XIV Storage System configuration, from six modules to 15 modules, has different performance capabilities. We call this *performance capacity*.

Performance capacity is different from the data storage capacity associated with each XIV Storage System configuration. If the data storage requirement indicates that a 12-module XIV Storage System is needed to store all the data and that a 15-module system is required to satisfy the performance requirements of the applications, then a 15-module system is required.

In addition to the number of modules, another way to properly balance the needs of both capacity and performance in tandem is to appropriately select the hard disk drive (HDD) capacity of the XIV Storage System configuration. However, for purposes of planning for future storage needs, it is also important to remember that the IBM XIV systems support homogeneous drive configurations only. Therefore, ensure that the initial selection of drive capacity is viable in terms of forecasting capacity growth requirements.

Note: Gen3 (Model 114) systems running Version 11.1.1 or later of the XIV system software support 1 TB of capacity per spindle. Using smaller drives maximizes the ratio of dynamic random access memory (DRAM) and flash cache to disk capacity (also called the *backstore*). Using smaller drives also increases the potential maximum viable aggregate workload access density (the ratio of IOPS to each unit of consumed capacity). In performance terms, the result is that the aggregate potential working set size that can be achieved with 100% capacity use is greatest when 1 TB drives are selected. In addition to the impact to caching, also consider that decreasing the drive size obviously decreases the maximum potential amount of data that can be accessed by each physical HDD arm. This means that the viable achievable physical disk IOPS per unit of capacity is maximized.

Performance requirements can be determined in several ways, including the following methods:

- Benchmark testing
- Application vendor specifications (based on benchmark testing)
- Existing I/O performance characteristics on existing storage being replaced by XIV Storage System

It is the client's responsibility to provide performance requirements. However, IBM licenses a performance modeling and sizing tool called *Disk Magic* that aligns client performance requirements with both the performance capacity and data storage capacity of an XIV Storage System configuration. Disk Magic is licensed to IBM and IBM Business Partners.

6.2.2 Number of logical unit numbers

When planning an XIV Storage System logical configuration, consider how many logical unit numbers (LUNs) to configure per host or per application.

With an XIV Storage System, there is no need to create many LUNs. As a matter of fact, when planning to deploy applications on an XIV Storage System, the primary design principle should be *fewer LUNs and larger LUNs*.

Every XIV Storage System LUN, regardless of size, is already, architecturally, automatically configured within the XIV Storage System for optimum performance. A planning item to any application migration to the XIV Storage System is the significant consolidation of the number of LUNs. Small quantities of volumes and LUNs are simpler to manage. Using fewer LUNs compared to a traditional storage subsystem improves read cache hit rates and space use because there are fewer "orphaned" physical capacity allocations.

However, if you must separate logs from data, or need to separate LUNs to create a disaster recovery strategy that includes snapshots, consistency groups, and volume replication, you might need to use more volumes. In addition, certain host platforms still require multiple volumes. The best approach is to examine the applications first and make a layout strategy that meets the application needs.

One of the most important considerations for the number of LUNs required for optimum performance in an XIV Storage System environment is application I/O threads. As a rule of thumb, if the application must use multiple LUNs to allocate or create multiple threads to handle the I/O, then use multiple LUNs. However, most modern enterprise applications are sophisticated enough to define multiple I/O threads independent of the number of LUNs, or the number of LUNs has no effect on application threads. In this case, there is no compelling reason to have multiple LUNs.

Multiple LUNs might be needed to use queues on the host HBAs or I/O buffers on the host operating system.

Another important consideration for the number of LUNs in an XIV environment has to do with parallelism of I/Os across the XIV interface modules. Using the multipathing design illustrated in Figure 6-3 as an example, consider one application server with one LUN:

- Each host is equipped with dual HBAs. Each HBA (or HBA port) is connected to one of two FC switches.
- Each of the FC switches has a connection to three separate interface modules.
- Each LUN has six paths.

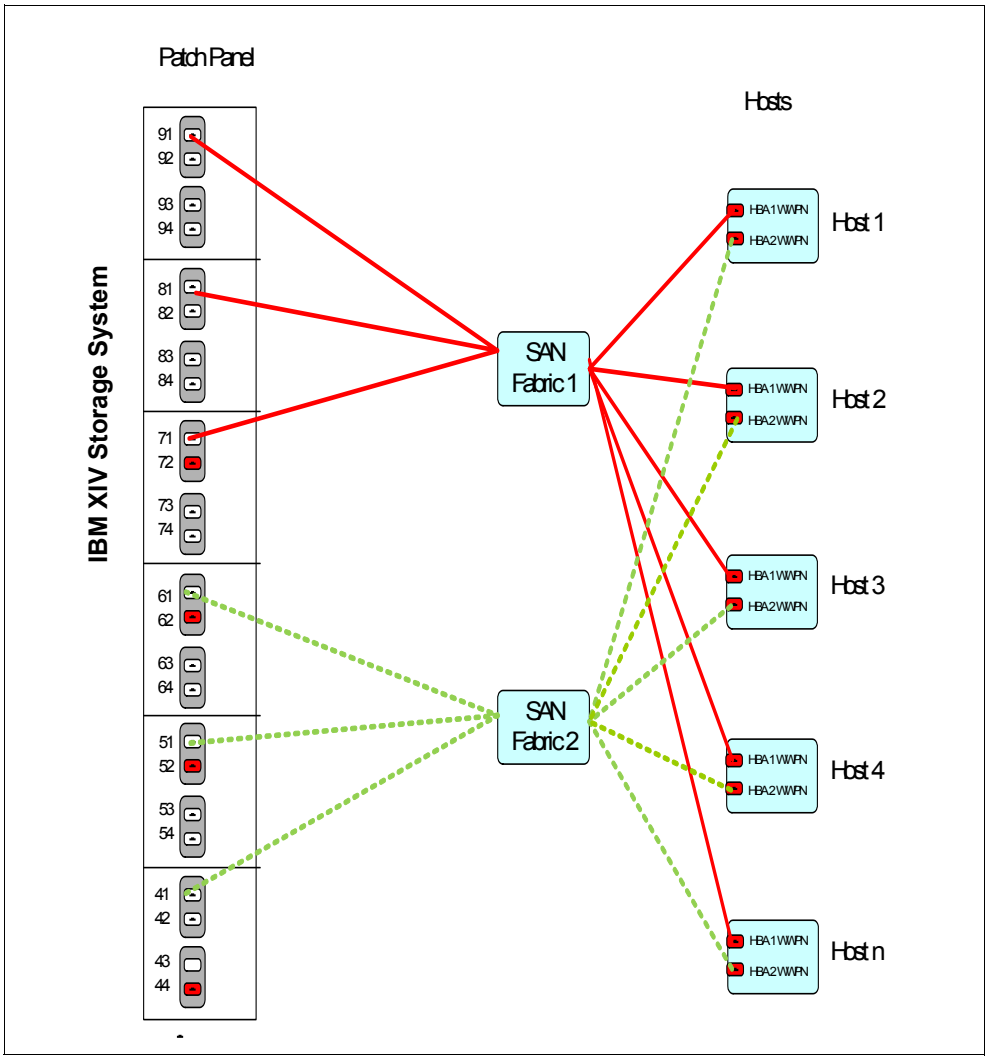


Figure 6-3 Six paths per LUN is the best overall multipathing configuration

This application server uses six paths in a round-robin fashion to perform I/O to this LUN. Each XIV Storage System interface module is used, one at a time, to handle this I/O. The net effect is a good balance between the interface modules. But in this example, the interface modules are being used only one at a time. For optimum performance, it is best to use all of the interface modules all of the time.

This task can be accomplished by using six host servers with one XIV Storage System LUN or by one host server with six LUNs. Because most computing environments use multiple

application servers, having one LUN per server is fine. But when the XIV is dedicated to one host system, consider matching the number of LUNs to the number of multipathing paths.

This consideration of the number of hosts and LUNs is closely linked to queue depth, as described in “Host bus adapter and disk queue depth” on page 299.

Compared to other storage systems, the XIV Storage System configuration typically ends up with a lower overall LUN count because the XIV Storage System architecture removes the limitations of traditional RAID arrays.

Older storage architectures often require small volumes (such as meta-volumes) to use more drive spindles. Older architectures also require many LUNs to employ striping techniques to engage enough disk spindles to handle high I/O requirements. XIV Storage System is optimized to use all drive spindles and eliminates the inconvenience of managing many volumes.

6.2.3 Multipathing considerations

XIV Storage System optimum performance is realized through the maximum use of the grid architecture. This use is the most important consideration to achieve maximum XIV Storage System performance.

Important: When planning host connections to XIV Storage System, it is important to ensure that all XIV Storage System interface modules are used.

For a description of multipathing, see Chapter 1, “Host connectivity,” in *XIV Storage System: Host Attachment and Interoperability*, SG24-7904. For the purposes of this book, we describe only Fibre Channel (FC) connectivity, although the concepts are the same for both the 1 Gbps or the 10 Gbps iSCSI host connectivity.

The main multipathing goal, from a performance perspective, is for the host connectivity to create a balance of the I/O workload across all of the resources in the XIV Storage System. The best way to achieve this balance is by distributing the host physical connections evenly across *all* of the interface modules. Providing host I/O access to every interface module from every host HBA has the following advantages:

- Uses the most XIV Storage System cache
- Uses the most XIV Storage System processor capability to handle I/O
- Fully uses the XIV Storage System grid architecture
- Minimizes the impact of a host interface hardware failure

There are two main multipathing techniques. The important point about both multipathing configurations is that each host engages the I/O services of each XIV Storage System interface module.

Experience has shown that the first multipathing configuration in Figure 6-3 on page 296 is the best overall general-purpose configuration. Host multipathing reliability during path error recovery in certain operating systems is complicated by increasing numbers of paths per LUN. Certainly, for host systems with two HBAs, the six paths per LUN method is the best way to go.

The second multipathing configuration, shown in Figure 6-4 on page 298, is more appropriate for benchmarking and higher-performance host systems with the highest I/O requirements. The primary difference is the host’s ability to handle the higher number of paths per LUN. It is *not* a good practice to use this configuration for most production applications, primarily

because of the complexity that the added paths per LUN imposes on the host multipathing software.

Notice these features of the configuration example in Figure 6-4:

- Each host is equipped with dual HBAs. Each HBA (or HBA port) is connected to one of two FC switches.
- Each of the FC switches has a connection to a separate FC port of each of the six interface modules.
- Each LUN has 12 paths.
- Each host HBA is physically connected to all six of the XIV Storage System interface modules. This setup provides the ability for each HBA to involve the maximum available I/O capabilities.

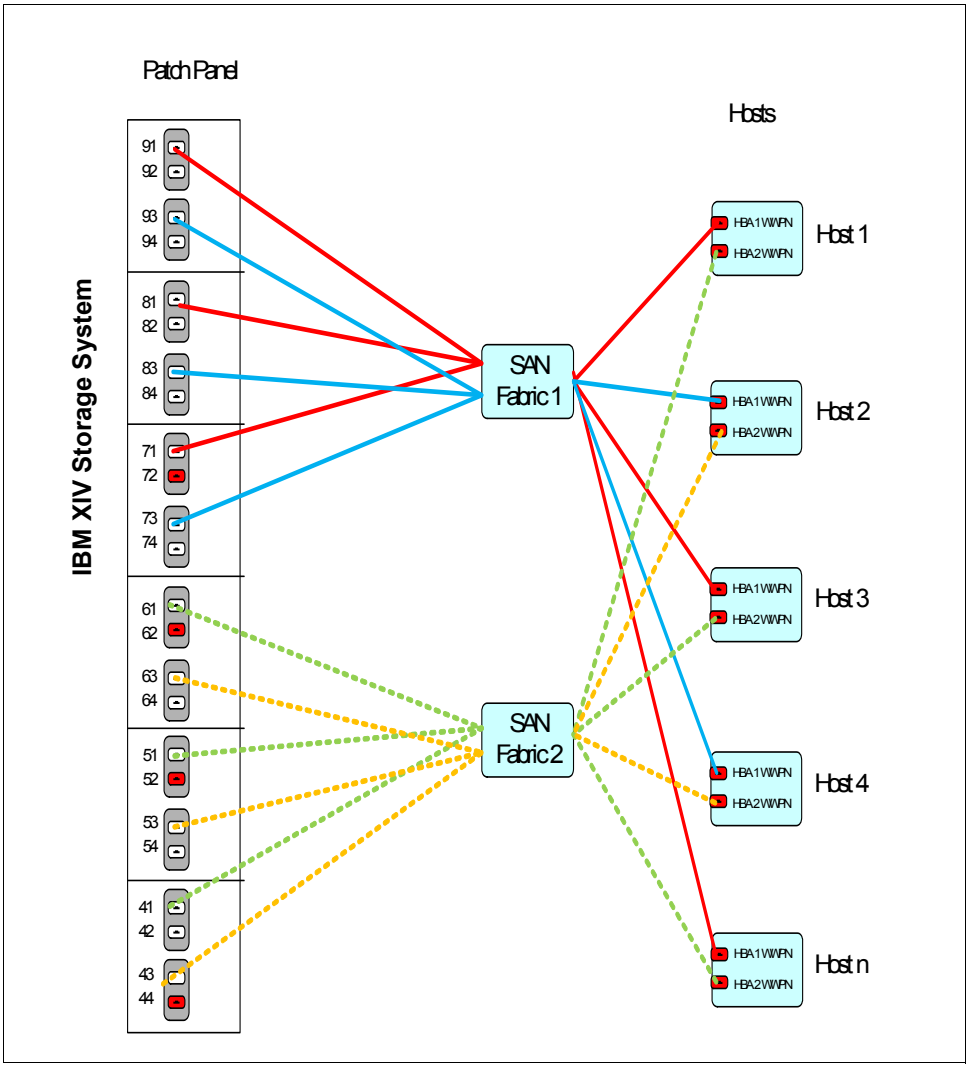


Figure 6-4 Multipath configuration for larger hosts

For XIV Storage System, each interface module has two 2-port Fibre Channel adapters. It is a good practice that each zone is physically connected to one port on each XIV Storage System Fibre Channel adapter. For example, switch A's zone can be connected to port 1 on the interface modules and switch B's zone can be connected to port 3 on the Interface Modules.

6.2.4 Host considerations

The XIV Storage System provides maximum performance when the application hosts can engage all the components of the XIV Storage System grid architecture with parallel I/O requests to the LUNs. Ideally, the host workload can be tailored to use multiple threads to simultaneously use multiple paths and access data on multiple volumes.

The following considerations can significantly affect XIV Storage System performance:

- Application threads
- HBA and disk queue depth
- Logical Volume Manager (LVM) striping
- Operating system tunables

Application threads

The XIV Storage System grid architecture excels at and experiences peak performance when applications employ multiple threads to handle the parallel execution of I/Os. All modern commercial applications employ multiple I/O threads. The number of threads are often tunable. From an XIV Storage System performance perspective, it is best to use the maximum number of threads possible without having a negative impact performance impact on the application within the host server.

For single-threaded applications, if there is no compelling reason to run them in a serial fashion, run as many of them in parallel as makes sense.

Host bus adapter and disk queue depth

Queue depth is another example of how the XIV Storage System architecture performs best with real-world client production workloads.

Queue depth is an important host HBA setting because it essentially controls how much data is allowed to be “in flight” onto the SAN from the host HBA. A queue depth of one requires that each I/O request is completed before another is started. A queue depth greater than one indicates that multiple host I/O requests can be waiting for responses from the storage system. So, the higher the queue depth, the more parallel I/O goes to the XIV Storage System.

The disk queue depth is an important OS setting as well that controls how much data is allowed to be “in flight” for a certain XIV Storage System volume to the HBA. The disk queue depth depends on the number of XIV Storage System volumes attached to this host from one XIV Storage System and the HBA queue depth. For example, if you have a host with just one XIV Storage System volume attached and two HBAs with an HBA queue depth of 64, you must configure a disk queue depth of 128 for this XIV Storage System volume to be able to fully use the queue of the HBAs.

The XIV Storage System architecture eliminates the common storage concept of a large central cache. Instead, each component (module) in the XIV Storage System grid has its own dedicated cache. The XIV Storage System algorithms that stage data between disk and cache work most efficiently when multiple I/O requests are coming in parallel, which is when the host queue depth becomes an important factor in maximizing XIV Storage System I/O performance. It is usually best to configure large host queue depths to ensure that you use the parallelism of the XIV Storage System architecture. A good practice is starting with a queue depth of 64 per HBA to ensure use of the XIV Storage System parallel architecture.

Tip: A queue depth of 64 is the best host HBA queue depth to start with for planning purposes.

Figure 6-5 shows a queue depth comparison for a database I/O workload (70 percent reads, 30 percent writes, 8 K block size, and DBO = Database Open).

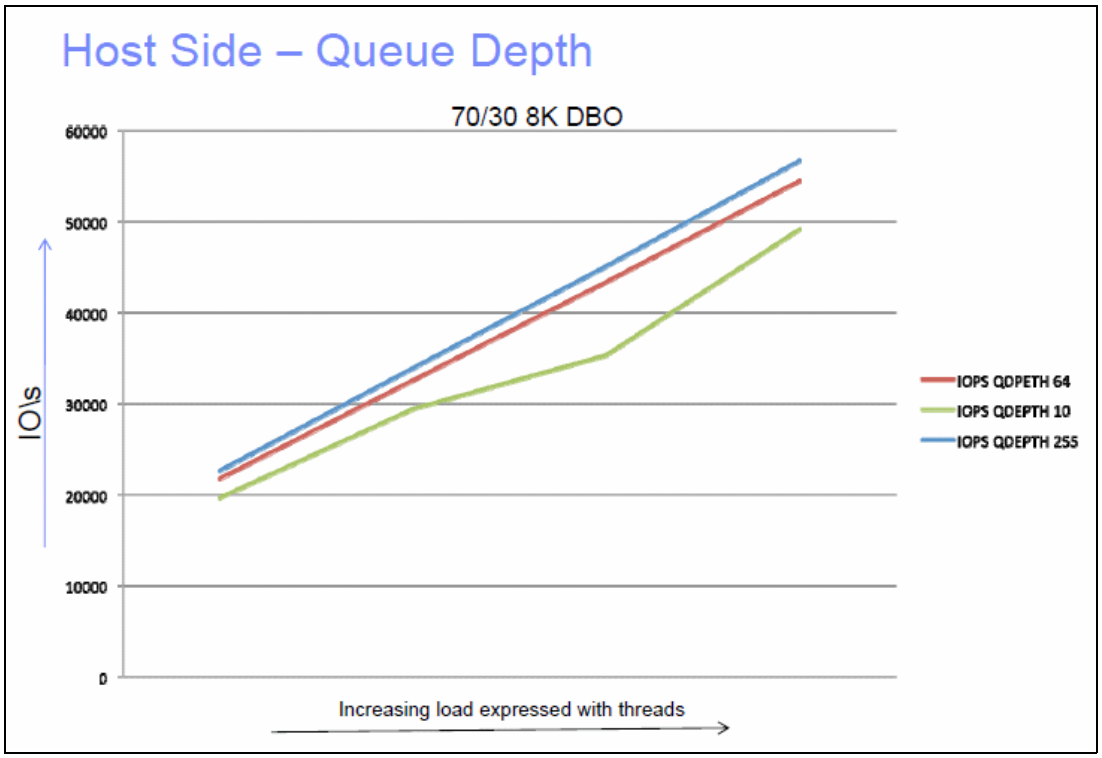


Figure 6-5 Host side queue depth comparison

Disclaimer: Performance numbers in this example are valid only for this special test conducted at an IBM lab. The numbers do not describe the general capabilities of an XIV Storage System as you might observe them in your environment.

The green line in the graph is the performance of a host with a queue depth set to 10. Where this line changes slope near 35,000 IOPS, you can see the red line at around 44,000 IOPS. The red line represents the same host configured with HBA queue depth settings of 64. The difference here at the widest margin is around a 25% boost in IOPS just by changing the host queue depth setting from 10 to 64.

The blue line shows the queue depth, which is set to 256. Having the queue depth this high does not provide much of an advantage or disadvantage.

Higher queue depth in general yields better performance with XIV Storage System. It is also important to consider the limitations per port on the XIV Storage System side.

For example, each HBA port on an XIV Storage System interface module sustains up to 1400 concurrent I/Os (except for model A14 on port 3 when port 4 is defined as an initiator, in which case port 3 is set to sustain up to 1000 concurrent I/Os). With a queue depth of 64 per-host-port, one XIV Storage System port is limited to 21 concurrent host ports, assuming that each host fills up the entire 64 depth queue for each request. The point here is that for larger configurations (more than 20), it is important to include queue depth in your production planning.

Logical Volume Manager striping

In general, the XIV Storage System architecture does not require the use of host-based Logical Volume Manager (LVM) striping techniques to improve I/O performance because each XIV Storage System volume is already striped across all the drives in the XIV Storage System. Generally, it is better not to use any LVM striping or any other method that stripes the data (such as ASM on Oracle) unless it is required for specific reasons. If LVM striping is being considered, test the application with and without striping.

Normally, operating system striping techniques have a small negative impact on XIV Storage System performance. For example, LVM striping might adversely affect XIV Storage System caching algorithms.

Consider a large sequential read. Normally, an XIV Storage System aggressively prefetches this data into cache. But if this data exists on multiple LUNs because of LVM striping, as soon as the operating system calls for data from a different LUN, the sequential nature of the read is interrupted and must be detected again by XIV Storage System on the different LUN. XIV Storage System is not unique in this regard, because all caching storage systems experience this situation.

As mentioned earlier, there are reasons why LVM striping has performance benefits from a host perspective. An interesting example is with certain applications that are better able to use the performance capabilities of the host operating system by using LVM striping. These efficiencies are realized on the host kernel buffer use. This application was fully tested without LVM striping and with LVM striping using different stripe sizes. The net result was a 25% overall performance improvement using host LVM striping.

Operating system tunables

It is common to experience significant performance improvements by tuning the I/O characteristics of the operating system. Most of these characteristics are kernel I/O buffers, memory buffers, or file system tunables. The specifics for each operating system are complex and outside the scope of this book. This activity is a common one that cannot be overlooked during migration planning. You can find more information about this topic in *XIV Storage System: Host Attachment and Interoperability*, SG24-7904.

6.2.5 Quality of service

We previously introduced the quality of service (QoS) feature. QoS allows the XIV Storage System to deliver multiple service levels to hosts connected to the same system.

The QoS feature is intended to enhance performance of critical business applications that run concurrently with noncritical applications. As the XIV Storage System disk and cache are shared among all applications, and all hosts are attached to the same resources, division of these resources among both critical and noncritical applications might have an unintended adverse performance on critical applications. In response to this issue, limiting the rate of noncritical IOPS and bandwidth by specifying and then enforcing limits on the maximum amount of low priority IOPS and bandwidth on a host basis is ideal.

As a result, the QoS feature in the XIV Storage System enables better performance for the critical host applications that run concurrently with the noncritical host applications on the same XIV Storage System.

We explained in 4.7, “QoS feature” on page 170 how to set QoS, by defining performance classes in terms of IOPS and bandwidth limitation and then assigning specific hosts to a particular performance class. Each host can be assigned only a single performance class at a time. However, there is no limit in the number of hosts within a specified class.

Limitation by bandwidth

The interface modules enforce the configured limitations. The intended limitation value depends on the number of interface modules being used by the hosts within the same performance class. The maximum rate value specified is multiplied by the number of interface modules installed in the XIV system to determine the rate for the class.

For example, a noncritical host is connected to all six interface modules:

- If the application administrator intends a 300 MBps limit for that host, the administrator user must set the QoS bandwidth limit for that host to 300 and the Bandwidth Limit per Interface is automatically set to 50.
- With six interface modules, the enforcement is 50 MBps per interface module, thereby limiting the host to an aggregate bandwidth of 300 MBps (50 MBps x 6 Modules = 300 MBps). If only two interface modules were used, the limit for the host is 100 MBps (50 MBps x 2 Modules = 100 MBps).
- If the host has connections to only two interface modules in a full six-interface module system, the actual host bandwidth limitation is only 100 MBps with this performance class setting (50 MBps x 2 Modules = 100 MBps). Therefore, if the user intends to have a 300 MBps bandwidth limitation with two interface modules connected in a full six interface module system, the Bandwidth Limit per Interface is 150 MBps and the Bandwidth Limit must be set to 900.

Limitation by input/output operations per second

Again, the interface modules enforce the limitations.

If the intent is to set the limitation at 10,000 IOPS for a specified host in a six-interface module configuration, the IOPS limit must be set to 10,000 and the enforcement is 1,666 (10,000/6) for each interface module.

If the host is attached to just two interface modules in a full six-interface module system, the host IOPS limitation would be only 3332 with this performance class setting (1666 IOPS x 2 interface modules = 3332 IOPS).

If the intent is to have a 10 K IOPS limitation for a host connected to only two interface modules in a full six-interface module system, “IOPS Limit Per Interface” must be set to 5000 or “IOPS Limit” for the performance class needs to be set to 30,000. Users must consider these interface module multiplication factors to properly meet expected limitations.

Quality of service illustration

To illustrate the QoS effect, we used it with two OLTP applications (one designated as critical, the other noncritical) running on the XIV system. They were set to run at about 30 K IOPS total. We repeated the test with three different levels of QoS limitations (1 K, 5 K, and 10 K). In each test, both applications start simultaneously. The noncritical application uses the QoS after applications first run concurrently for 5 to 10 minutes. Within one minute after enforcing QoS, the critical application shows performance improvements.

The critical application in Figure 6-6 shows an IOPS improvement from 15 K to 27 K (76%) by limiting noncritical OLTP host application IOPS to 1 K. The critical OLTP host application IOPS also improved to 53% and 23%, respectively, after limiting noncritical host applications to 5 K and 10 K.

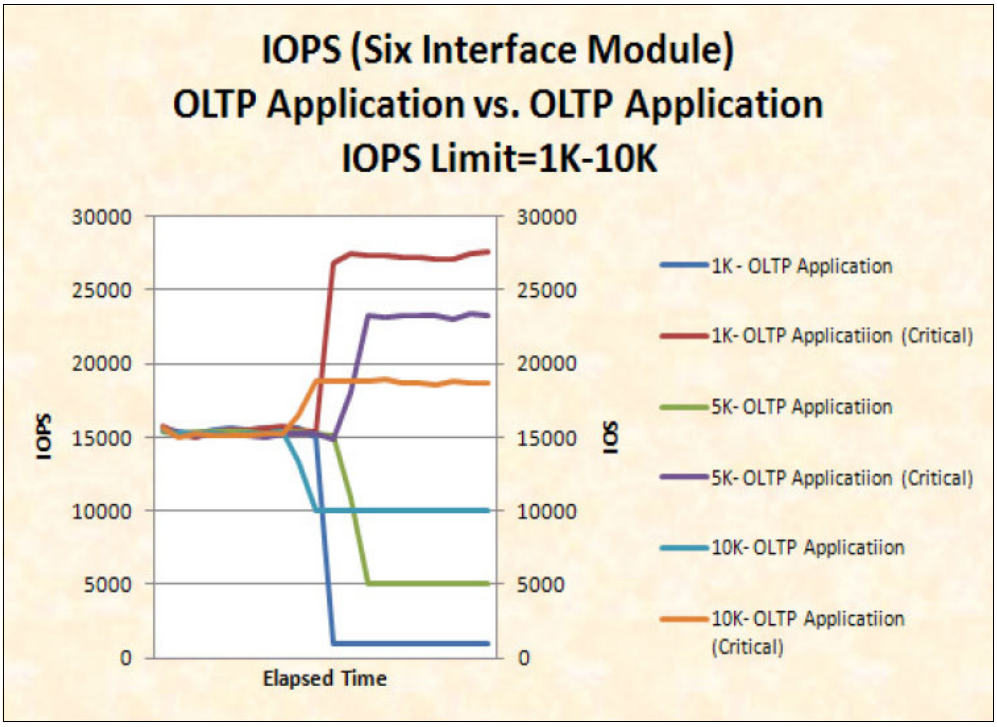


Figure 6-6 QoS illustration IOPS

The average I/O response time in both shows similar behavior. The largest response time (RT) was at the 1 K IOPS limit, as shown in Figure 6-7.

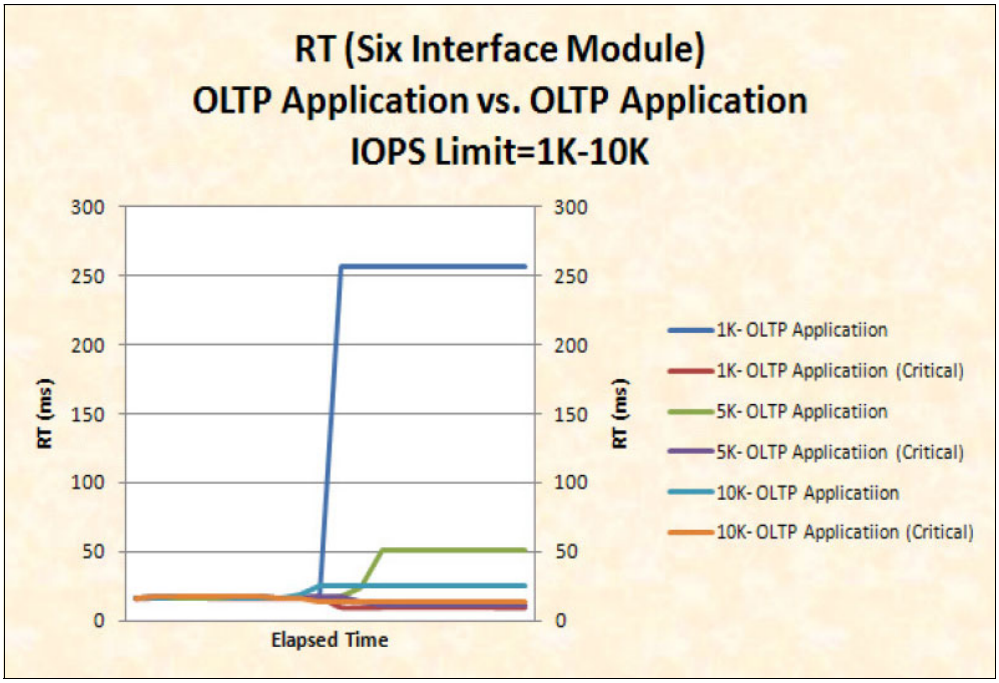


Figure 6-7 QoS illustration response time (RT)

6.3 Performance monitoring

During normal operation, the XIV Storage System constantly gathers statistical information. The data can then be processed using the XIV Storage Management GUI, XIV Storage System Command-Line Interface (XCLI), or IBM Tivoli Storage Productivity Center for Disk. This section introduces the techniques for processing the statistics data.

6.3.1 Using the XIV Storage Management GUI

The XIV Storage Management GUI provides a mechanism to gather statistics data. For a description of setting up and using the XIV Storage Management GUI, see Chapter 4, “IBM XIV Storage Management software” on page 111. When working with the statistical information, the XIV Storage System collects and maintains the performance data internally. As the performance data ages, it is consolidated to save space according to the following rules:

- Data is kept for 365 days.
- Detailed host and volume data is kept for 30 days.

By selecting specific filters, the requested data is mined and displayed. This section describes the functions of the XVI Storage Management GUI and how to retrieve the required data.

The first item to note is that the current IOPS for the system is always displayed in the bottom center of the window. This feature provides simple access to the current load on the system. Figure 6-8 illustrates the XIV Storage Management GUI and the IOPS display. It also shows how to start the statistics monitor.

Figure 6-8 Starting the statistics monitor on the GUI

Select **Statistics** from the Monitor menu, as shown in Figure 6-8 on page 304, to open the monitor default window that is shown in Figure 6-9. Figure 6-9 shows the system IOPS.

Figure 6-9 Default statistics monitor view

In Figure 6-9, the X-axis of the graph represents the time and can vary from minutes to months. The Y-axis of the graph is the measurement that is selected. The default measurement is IOPS.

The statistics monitor can also illustrate latency and bandwidth in the same graph or in the multigraph view.

The other options in the statistics monitor act as filters for separating data. These filters are separated by the type of transaction (reads or writes), cache properties (hits compared to misses) and cache memory hits or flash (SSD) cache hits, or the transfer size of I/O as seen by the XIV Storage System. The filter pane has been updated to reflect support for SSD Caching and to filter between read I/Os from the main cache (Mem Hit) and read I/Os from the extended SSD cache.

Figure 6-10 shows the filter pane.

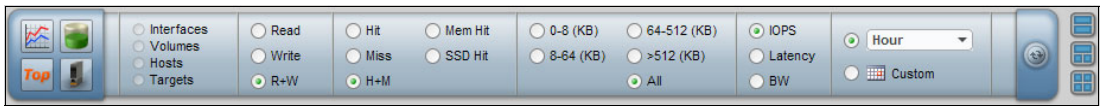


Figure 6-10 Filter pane for the statistics monitor

The filter pane allows you to select multiple items within a specific filter, for example, if you want to see reads and writes separated on the graph. By holding down `Ctrl` on the keyboard and selecting the read option and then the write option, you can see both items on the graph.

As shown in Figure 6-11, one of the lines represents the read IOPS and the other line represents the write IOPS. On the GUI, these lines are drawn in separate colors to differentiate the metrics. The other most popular speed metric is bandwidth (BW), measured in MBps.

This selection process can be performed on the other filter items as well.

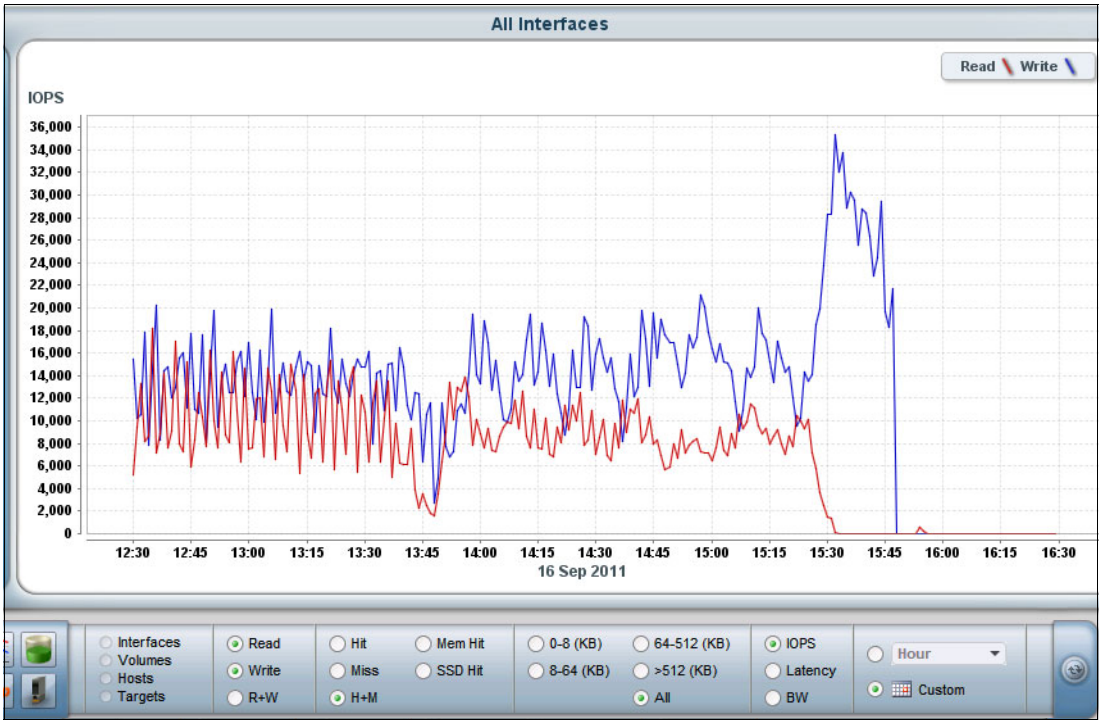


Figure 6-11 Multiple filter selection

One of the most meaningful performance metrics is latency, which is measured in milliseconds. Figure 6-12 on page 307 shows the read and write latency for the entire XIV system.

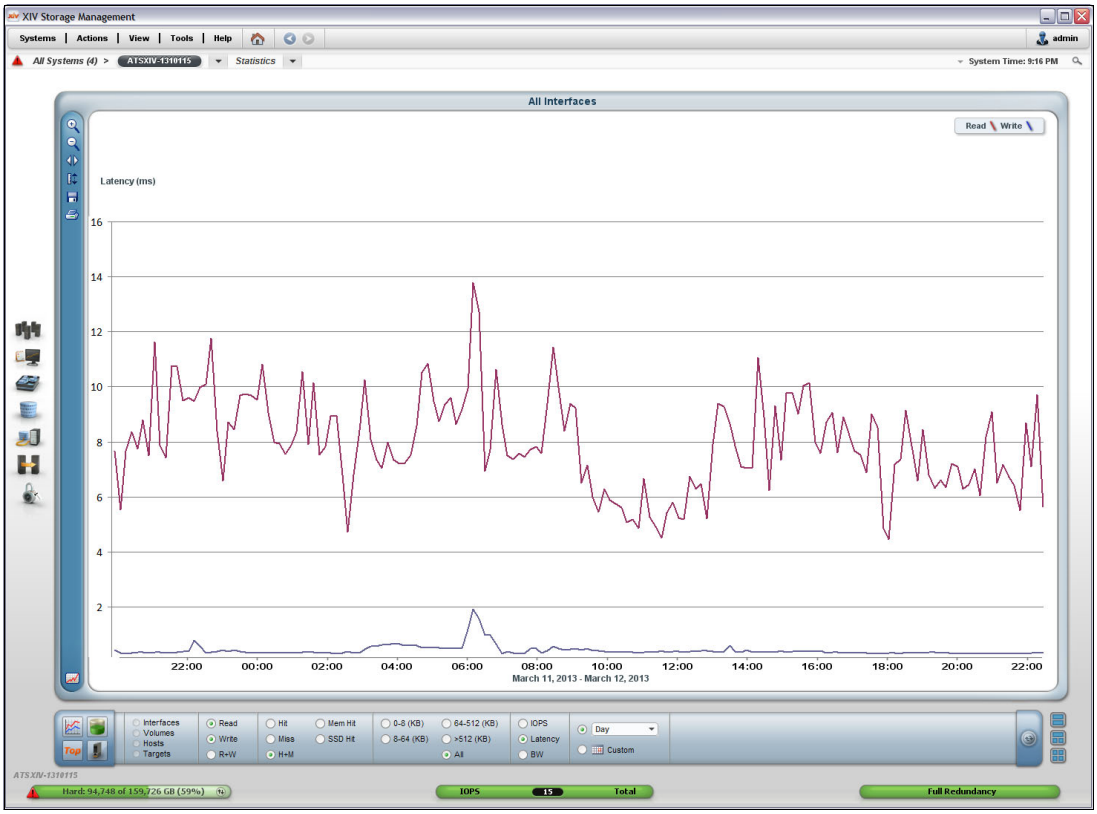


Figure 6-12 Latency

Latency (and all the other metrics, for that matter) can also be shown for individual or multiple Fibre Channel or iSCSI interfaces, volumes, or hosts. For more information about latency, see “Performance analysis” on page 321.

Another popular view of performance is the percentage of read hits as shown in Figure 6-13 on page 308. Read hits are the total read requests that are satisfied by reading the data from the XIV system cache. Also shown are the cache hits satisfied from the DRAM memory cache and the SSD cache. The read cache misses shown are the read requests that were satisfied by retrieving the data from the disk drives. Again, there is more information about latency in “Performance analysis” on page 321.

Figure 6-13 Read hits

In certain cases, the user needs to see multiple graphs at one time. On the right side of the filter pane, there is a selection to add graphs (see Figure 6-10 on page 305). Up to four graphs are managed by the GUI. Each graph is independent and can have separate filters. Each of the multiple graphs can be unlocked to show a different time duration than the others.

Figure 6-14 on page 309 shows this multigraph concept. The top graph is the IOPS for the day with separated reads and writes. The second graph shows the bandwidth for a few minutes with separated reads and writes, which provides quick and easy access to multiple views of the performance metrics.

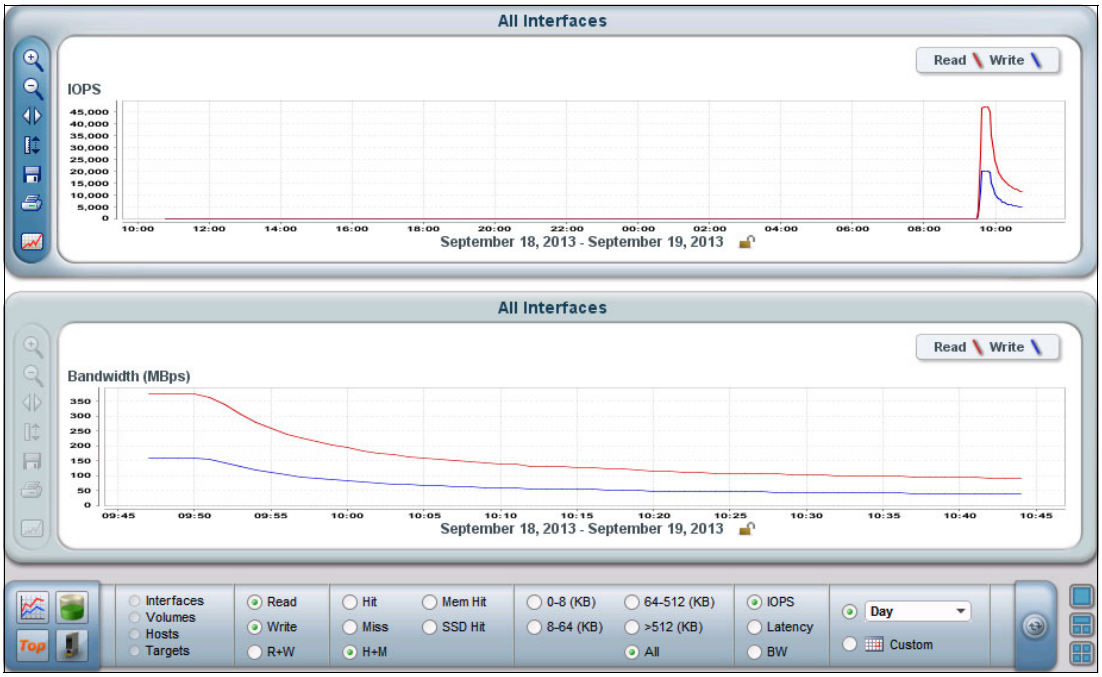


Figure 6-14 Multiple graphs using the XIV Storage Management GUI

There are various additional filters available, such as filtering by host, volumes, interfaces, or targets. These items are defined on the left side of the filter pane. When clicking one of these filters, a window opens. Highlight the item, or select a maximum of four by using the Ctrl key, to be filtered, and then select **click to select**. This action moves the highlighted item to the lower half of the window. To generate the graph, you must click the green check mark on the lower right side of the window. Your new graph is generated with the name of the filter at the top of the graph. See Figure 6-15 for an example of this filter.

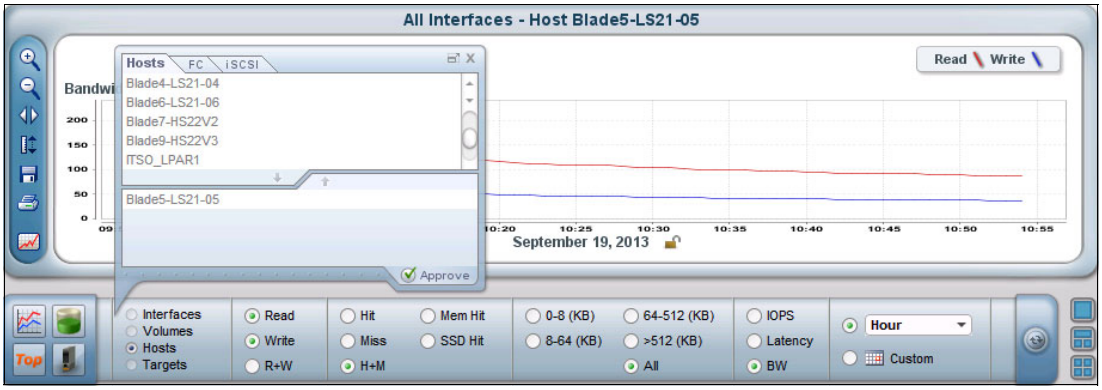


Figure 6-15 Example of a host filter

On the left side of the chart in the blue bar, there are various tools to assist you in managing the data.

Figure 6-16 shows the chart toolbar in more detail.

Figure 6-16 Chart toolbar

The top two tools (magnifying glasses) zoom in and out for the chart, and the second set of two tools adjusts the X-axis and the Y-axis for the chart. Finally, the bottom two tools allow you to export the data to a comma-separated file or print the chart to a printer.

On the right end of the filter pane is a time duration selector. These selections are helpful to quickly navigate to different times. In Figure 6-17, if “Hour” is selected, the time scale is reset to the last hour from the *current* time. Similarly, Day, Week, Month, and Year can be selected to easily see the performance statistics for these different time durations. Also, notice the refresh button to the right, which updates the graph with the latest information.

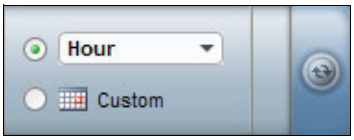


Figure 6-17 Time duration selector

Also, you can select **Custom** and a window opens where you can specify a certain date and time (Figure 6-18). After you specify the date and time, you can select the white box at the bottom of the window and a one-hour view of performance data ending at the selected date and time is displayed. To change the view from this specific date and time, you must use the chart toolbars in the upper left corner of the Statistics view.

Figure 6-18 Custom date and time selector

Another interesting view of XIV Storage System performance in the XIV Storage Management GUI can be found by selecting the icon of an XIV Storage System disk drive found in the lower left corner of the XIV Storage Management GUI, as shown in Figure 6-19.

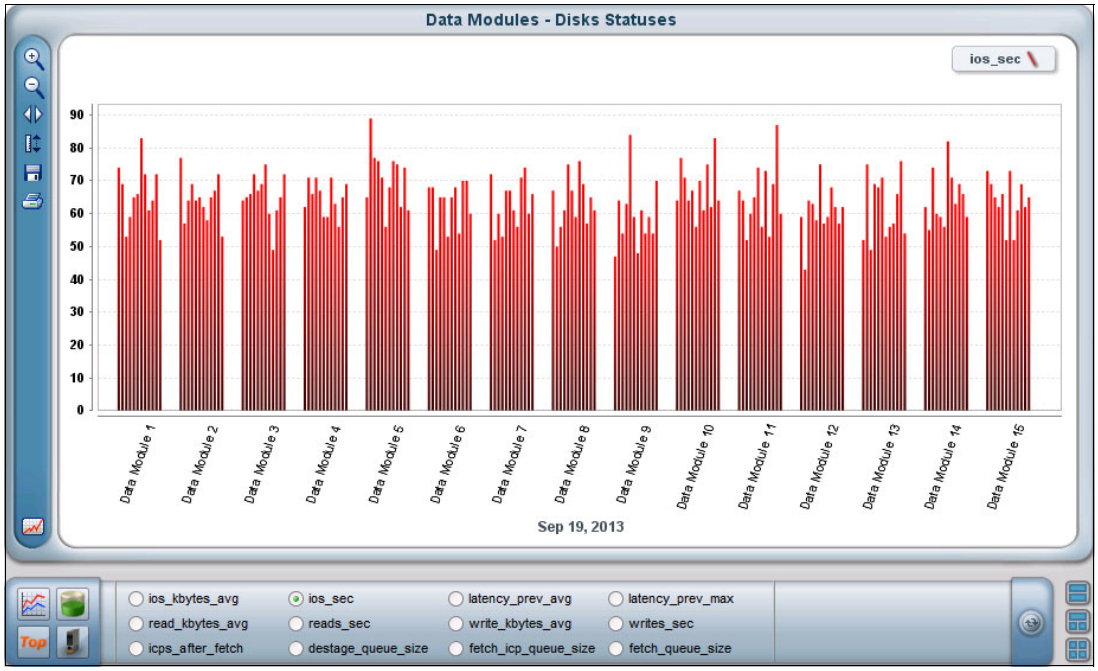


Figure 6-19 Disk performance view

This disk performance view shows a bar chart of every disk drive in the XIV Storage System configuration. There are several different performance metrics from which to choose. Here is an explanation of what each one means:

ios_kbytes_ave	Average I/O size per disk.
read_kbytes_ave	Average I/O read size per disk.
destage_icp_queue_size	Queue size of the pending Instant Copies (ICPs) we have for a disk. This metric is related to snapshot and write workload.
destage_queue_size	Number of write commands that are waiting in the queue for the disk. This metric increases as writes increase from cache to disk.
ios_sec	Number of IOPS for each disk, including IOPS in buffer queues. 100 is normal. 300 is considered high.
reads_sec	Number of read IOPS.
writes_sec	Number of write IOPS.
latency_prev_avg	Average latency.
write_kbytes_avg	Average I/O write size.
fetch_icp_queue_size	Queue size of the pending read ICP.
latency_prev_max	Max latency.
fetch_queue_size	Number of read commands that are waiting in the disk queue.

6.3.2 Using the XIV Top utility

In addition to the XIV Storage Management GUI performance statistics options, you can also use XIV Top for additional and more granular system performance statistics. The XIV Top application allows you to view and monitor performance information for defined volumes and hosts in real time.

There are multiple ways to access the XIV Top utility:

- To start XIV Top directly from the GUI desktop, right-click within the “All systems” view on the XIV Storage System that you want to monitor. See Figure 6-20.

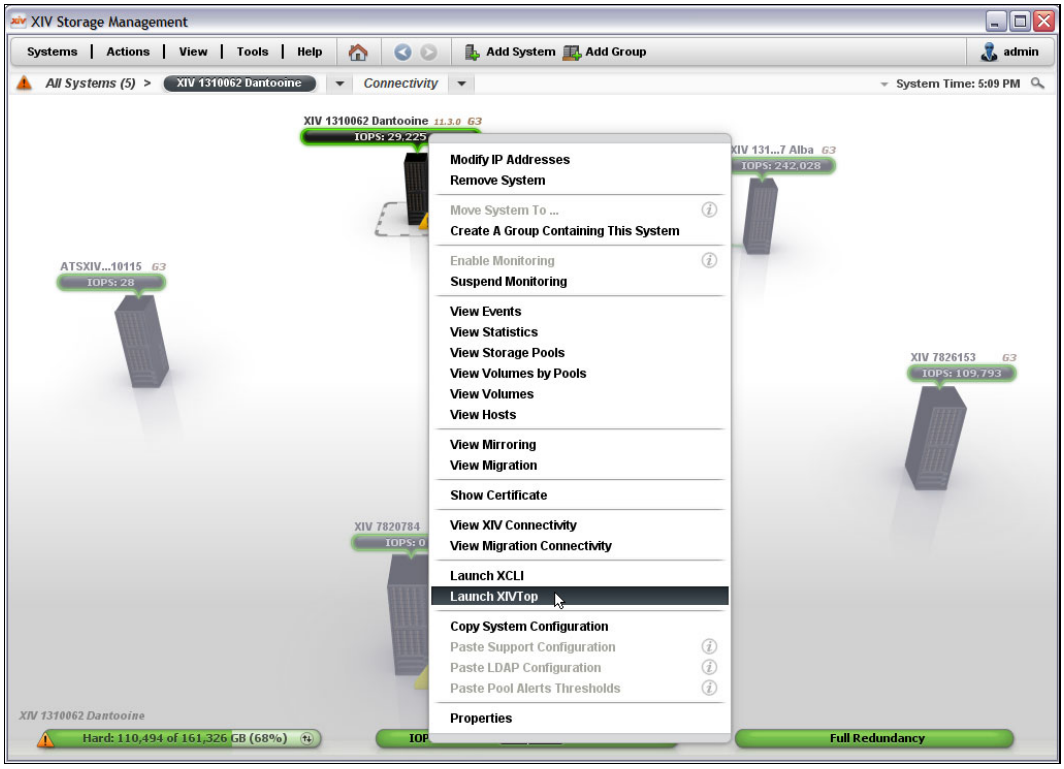


Figure 6-20 Starting XIV Top from the main GUI desktop

- Start XIV Top from the XIV Storage Management GUI desktop (Figure 6-21).

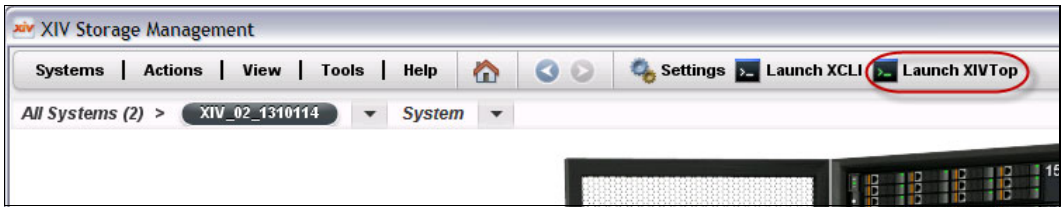


Figure 6-21 Start XIV Top from the XIV Storage Management GUI Desktop

- XIV Top can also be started within the statistic view from the lower left corner of the XIV Storage Management GUI Statistics Monitor Filter pane (see Figure 6-10 on page 305).

The XIV Top tool that is shown in Figure 6-22 quickly shows the volumes or hosts that are consuming the most system resources at any given time. Information can be sorted in various ways and exported to a file for use in spreadsheet applications.

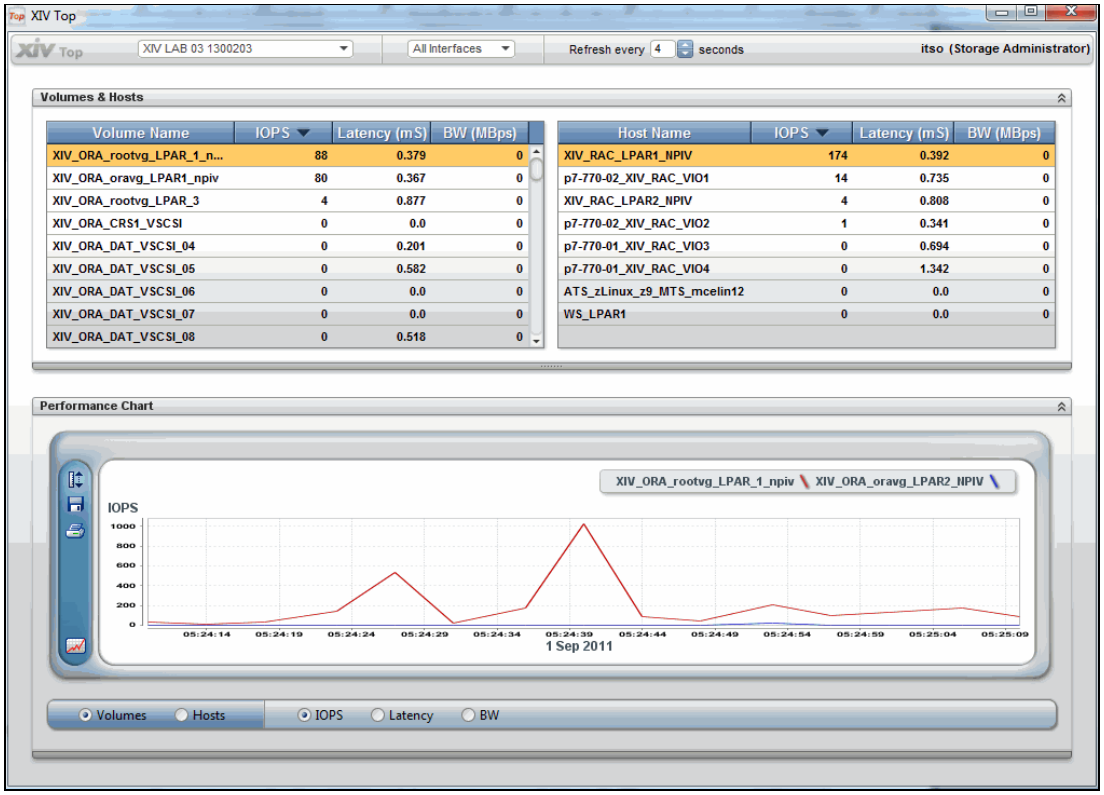
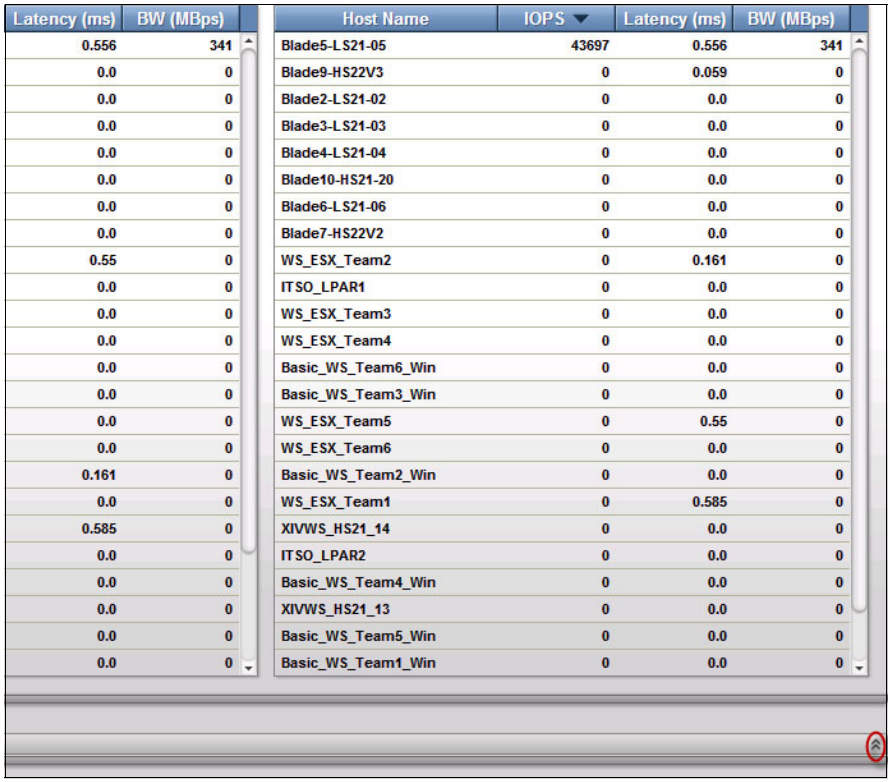


Figure 6-22 XIV Top monitoring tool

The initial XIV Top window is divided into two horizontal sections, the Volumes and Hosts section and the Performance Chart section. You can collapse or expand each section by clicking the arrow icon on the right side of the header of each section (Figure 6-23).



Latency (ms)	BW (MBps)	Host Name	IOPS	Latency (ms)	BW (MBps)
0.556	341	Blade5-LS21-05	43697	0.556	341
0.0	0	Blade9-HS22V3	0	0.059	0
0.0	0	Blade2-LS21-02	0	0.0	0
0.0	0	Blade3-LS21-03	0	0.0	0
0.0	0	Blade4-LS21-04	0	0.0	0
0.0	0	Blade10-HS21-20	0	0.0	0
0.0	0	Blade6-LS21-06	0	0.0	0
0.0	0	Blade7-HS22V2	0	0.0	0
0.55	0	WS_ESX_Team2	0	0.161	0
0.0	0	ITSO_LPAR1	0	0.0	0
0.0	0	WS_ESX_Team3	0	0.0	0
0.0	0	WS_ESX_Team4	0	0.0	0
0.0	0	Basic_WS_Team6_Win	0	0.0	0
0.0	0	Basic_WS_Team3_Win	0	0.0	0
0.0	0	WS_ESX_Team5	0	0.55	0
0.0	0	WS_ESX_Team6	0	0.0	0
0.161	0	Basic_WS_Team2_Win	0	0.0	0
0.0	0	WS_ESX_Team1	0	0.585	0
0.585	0	XIVWS_HS21_14	0	0.0	0
0.0	0	ITSO_LPAR2	0	0.0	0
0.0	0	Basic_WS_Team4_Win	0	0.0	0
0.0	0	XIVWS_HS21_13	0	0.0	0
0.0	0	Basic_WS_Team5_Win	0	0.0	0
0.0	0	Basic_WS_Team1_Win	0	0.0	0

Figure 6-23 XIV Top with Performance Chart section collapsed

Each column of the Volumes and Hosts section can be sorted in ascending or descending order by clicking the column and toggling the direction of the sort indicator (Figure 6-24 on page 315).

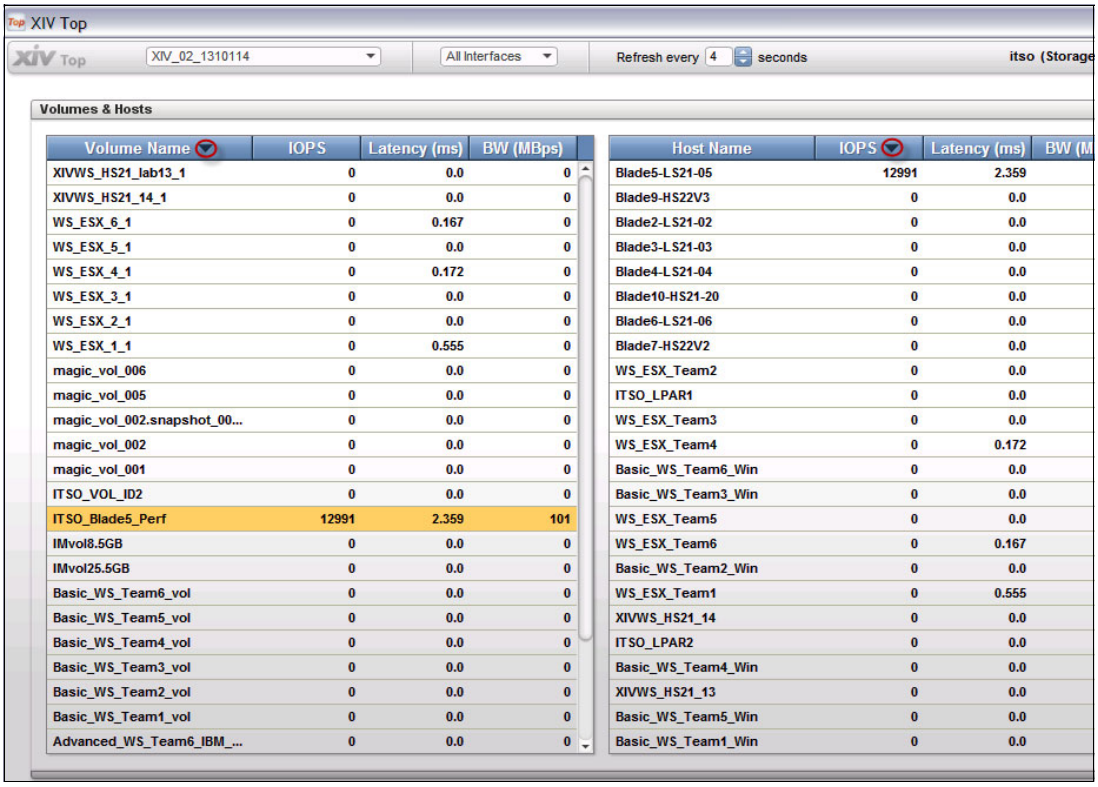


Figure 6-24 Sort columns in ascending or descending order

The IOPS, latency, or bandwidth for up to four volumes or four hosts can be viewed in the Performance Chart section. The controls to select the volumes, hosts, and options operate in much the same way as the Statistics tool from within the GUI (Figure 6-25).

Figure 6-25 Display performance information for up to four volumes or hosts with XIV Top

6.3.3 Using the XIV Storage System command-line interface

The second method to collect statistics is by using the XIV Storage System command-line interface (XCLI) operation. To access the XCLI Session, see Chapter 4, “IBM XIV Storage Management software” on page 111.

First, you must retrieve the system’s time. To retrieve the system’s time, issue **time_list**, and the system retrieves the current time. See Example 6-1 for an example of retrieving the XIV Storage System time.

Example 6-1 Retrieving the XIV Storage System time

```
>> time_list
Time       Date        Time Zone   Daylight Saving Time
11:45:42   2010-09-27  GMT         no
```

After the system time is obtained, the **statistics_get** command can be formatted and issued.

The **statistics_get** command requires various parameters to operate. The command requires that you enter a starting or ending time point, a count for the number of intervals to collect, the size of the interval, and the units related to that size. The time stamp is modified by the previous **time_list** command. Example 6-2 provides a description of the command.

Example 6-2 The statistics_get command format

```
statistics_get [ perf_class=perfClassName | host=HostName |
host_iscsi_name=initiatorName | host_fc_port=WWPN | target=RemoteTarget |
remote_fc_port=WWPN | remote_ipaddress=IPAddress | vol=VolName |
ipinterface=IPInterfaceName | local_fc_port=ComponentId ] < start=TimeStamp |
end=TimeStamp > [ module=ModuleNumber ] count=N interval=IntervalSize
resolution_unit=<minute|hour|day|week|month>
```

To further explain this command, assume that you want to collect 10 intervals, and each interval is for 1 minute. The point of interest occurred on 16 June 2009 roughly 15 minutes after 11:45:00. It is important to note that the **statistics_get** command allows you to gather the performance data from any time period.

The time stamp is formatted as YYYY-MM-DD:hh:mm:ss, where the YYYY represents a four-digit year, MM is the two-digit month, and DD is the two-digit day. After the date portion of the time stamp is specified, you specify the time, where hh is the hour, mm is the minute, and ss represents the seconds.

Example 6-3 shows a typical use of this command.

Example 6-3 The statistics_get command example

```
>> statistics_get end=2009-06-16.11:45:00 count=10 interval=1
resolution_unit=minute
```

Figure 6-26 shows sample output of the statistics. The output shown is a small portion of the data provided.

Time	Read Hit Medium - IOps	Read Hit Medium - Latency	Read Hit Medium - Throughput
2009-06-16 11:35:00	3418	2851	111332
2009-06-16 11:36:00	3583	2769	119298
2009-06-16 11:37:00	3673	2797	121695
2009-06-16 11:38:00	3430	3121	113067
2009-06-16 11:39:00	3548	3108	117711
2009-06-16 11:40:00	3274	2933	109533
2009-06-16 11:41:00	2917	3270	96250
2009-06-16 11:42:00	2651	3262	87067
2009-06-16 11:43:00	2292	3168	74384
2009-06-16 11:44:00	2045	3449	65100

>>

Figure 6-26 Output from statistics_get command

Extending this example, assume that you want to filter out a specific host defined in the XIV Storage System. By using the host filter in the command, you can specify for which host you want to see performance metrics, which allows you to refine the data that you are analyzing.

See Example 6-4 for an example of how to perform this operation. See Figure 6-27 for a sample of the output for the command.

Example 6-4 The statistics_get command using the host filter

```
>> statistics_get host=adams end=2009-06-16.11:45:00 count=10 interval=1
resolution_unit=minute
```

Time	Read Hit Medium - IOps	Read Hit Medium - Latency	Read Hit Medium - Throughput
2009-06-16 11:35:00	149	2092	4489
2009-06-16 11:36:00	136	2689	4219
2009-06-16 11:37:00	141	3016	3897
2009-06-16 11:38:00	293	2927	8068
2009-06-16 11:39:00	418	3030	12574
2009-06-16 11:40:00	391	3101	12518
2009-06-16 11:41:00	445	3554	13858
2009-06-16 11:42:00	518	3536	15748
2009-06-16 11:43:00	490	3243	14352
2009-06-16 11:44:00	370	3628	11531

>>

Figure 6-27 Output from the statistics_get command using the host filter

In addition to the filter just shown, the **statistics_get** command can filter iSCSI names, host worldwide port names (WWPNs), volume names, modules, and many more fields. As an additional example, assume that you want to see the workload on the system for a specific module. The module filter breaks out the performance on the specified module. Example 6-5 pulls the performance statistics for module 5 during the same time period of the previous examples. Figure 6-28 on page 318 shows the output.

Example 6-5 The statistics_get command using the module filter

```
>> statistics_get end=2009-06-16.11:45:00 module=5 count=10 interval=1
resolution_unit=minute
```

```
Time                 Read Hit Medium - IOps   Read Hit Medium - Latency   Read Hit Medium - Throughput
2009-06-16 11:35:00  354                      894                         9980
2009-06-16 11:36:00  165                      831                         4485
2009-06-16 11:37:00  159                      813                         4194
2009-06-16 11:38:00  159                      846                         4166
2009-06-16 11:39:00  127                      852                         3374
2009-06-16 11:40:00  98                       848                         2493
2009-06-16 11:41:00  115                      856                         3080
2009-06-16 11:42:00  115                      842                         3124
2009-06-16 11:43:00  48                       811                         1262
2009-06-16 11:44:00  192                      896                         5036
>>
```

Figure 6-28 Output from statistics_get command using the module filter

6.3.4 Tivoli Storage Productivity Center

Tivoli Storage Productivity Center contains enhanced performance metrics for the XIV Storage System. For more information, see the Tivoli Storage Productivity Center web page:

http://www.ibm.com/systems/storage/software/center/index.html

On the IBM Redbooks main page, search for “Tivoli Storage Productivity Center” to find several Redbooks publications about how to use the product:

http://ibm.com/redbooks

For more information that relates to the XIV Storage System, see 7.7, “Using Tivoli Storage Productivity Center” on page 390.

It is possible to see comprehensive XIV Storage System I/O performance characteristics from Tivoli Storage Productivity Center. The most effective use of this data is to extract this data from the Tivoli Storage Productivity Center database into comma-separated values (CSV) files and then import them into a spreadsheet. This data is useful for performance reporting, analysis, and archiving.

Tivoli Storage Productivity Center is a good way to evaluate XIV Storage System performance. The Tivoli Storage Productivity Center data that is most useful for XIV Storage System performance is contained in the following reports:

- By Storage Subsystem (overall statistics on the entire XIV Storage System)
- By Module/Node
- By Volume
- By Port

Figure 6-29 shows where this data is in the Tivoli Storage Productivity Center GUI.

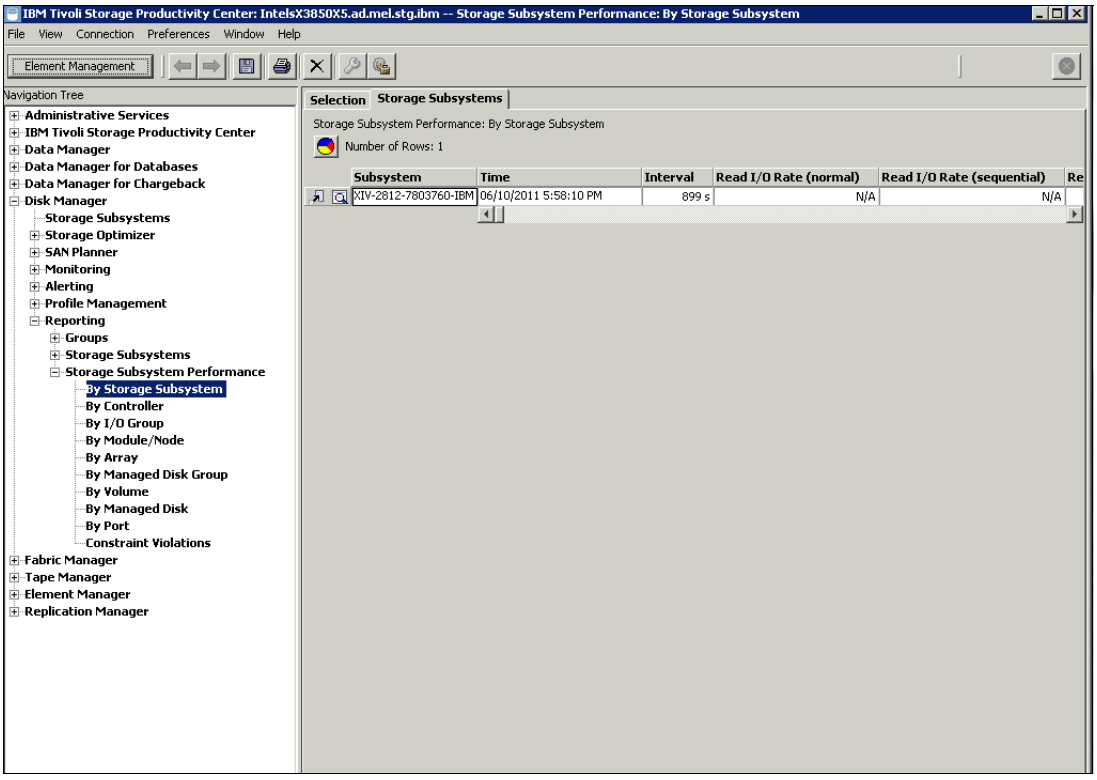


Figure 6-29 Tivoli Storage Productivity Center GUI

After loading this data into a spreadsheet, certain formatting is helpful to prepare the spreadsheet for the easiest use. See Figure 6-30 for the following formatting suggestions:

- Sort by date and time
- Freeze panes in cell E3
- Data filters in row 2

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	AM	AN	AO
1	Subsystem	Port	WWPN	Time	Interval	Port Send I/O Rate	Port Receive I/O Rate	Total Port I/O Rate	Port Send Data Rate	Port Receive Data Rate	Total Port Data Rate	Port Send Response Time	Port Receive Response Time	Total Port Response Time	Port Send Bandwidth Percentage	Port Receive Bandwidth Percentage	Overall Port Bandwidth Percentage
2																	
3	XIV-2812-7803760-IBM	1:FC_Port:4:1	500173800EB00140	27/09/11 10:13 PM	900	1748.19	877.48	2625.67	13.69	6.91	20.6006	12.4	1.1	8.6	3.34	1.68	2.51
4	XIV-2812-7803760-IBM	1:FC_Port:5:1	500173800EB00150	27/09/11 10:13 PM	900	1755.05	876.38	2631.43	13.71	6.92	20.6366	12.3	1.1	8.6	3.34	1.69	2.51
5	XIV-2812-7803760-IBM	1:FC_Port:6:1	500173800EB00160	27/09/11 10:13 PM	900	1748.74	876.65	2625.39	13.69	6.9	20.5969	12.4	1.1	8.6	3.34	1.68	2.51
6	XIV-2812-7803760-IBM	1:FC_Port:7:1	500173800EB00170	27/09/11 10:13 PM	900	1750.07	876.83	2626.91	13.7	6.93	20.6282	12.3	1	8.6	3.34	1.69	2.51
7	XIV-2812-7803760-IBM	1:FC_Port:8:1	500173800EB00180	27/09/11 10:13 PM	900	1749.06	878.22	2627.28	13.7	6.92	20.6191	12.3	1	8.5	3.34	1.68	2.51
8	XIV-2812-7803760-IBM	1:FC_Port:9:1	500173800EB00190	27/09/11 10:13 PM	900	1749.25	878.42	2627.66	13.69	6.95	20.6437	12.3	1	8.6	3.34	1.69	2.51
9	XIV-2812-7803760-IBM	1:FC_Port:4:1	500173800EB00140	27/09/11 10:28 PM	900	1736.98	871.87	2608.85	13.57	6.81	20.3832	12.4	1.1	8.6	3.31	1.66	2.48
10	XIV-2812-7803760-IBM	1:FC_Port:5:1	500173800EB00150	27/09/11 10:28 PM	900	1743.38	870.18	2613.56	13.61	6.79	20.3998	12.3	1.1	8.6	3.32	1.65	2.48
11	XIV-2812-7803760-IBM	1:FC_Port:6:1	500173800EB00160	27/09/11 10:28 PM	900	1738.53	871.46	2609.99	13.61	6.8	20.4071	12.4	1.1	8.6	3.32	1.66	2.49
12	XIV-2812-7803760-IBM	1:FC_Port:7:1	500173800EB00170	27/09/11 10:28 PM	900	1739.41	870.09	2609.5	13.6	6.79	20.393	12.3	1	8.6	3.32	1.65	2.48
13	XIV-2812-7803760-IBM	1:FC_Port:8:1	500173800EB00180	27/09/11 10:28 PM	900	1738.81	868.17	2606.98	13.6	6.78	20.3843	12.3	1	8.6	3.32	1.65	2.48
14	XIV-2812-7803760-IBM	1:FC_Port:9:1	500173800EB00190	27/09/11 10:28 PM	900	1738.21	873.36	2611.58	13.59	6.81	20.3991	12.3	1	8.5	3.31	1.66	2.48
15	XIV-2812-7803760-IBM	1:FC_Port:4:1	500173800EB00140	27/09/11 10:43 PM	900	1740.84	873.74	2614.58	13.63	6.81	20.4366	12.4	1.1	8.6	3.32	1.66	2.49
16	XIV-2812-7803760-IBM	1:FC_Port:5:1	500173800EB00150	27/09/11 10:43 PM	900	1745.24	871.34	2616.58	13.63	6.81	20.4323	12.4	1.1	8.6	3.32	1.66	2.49
17	XIV-2812-7803760-IBM	1:FC_Port:6:1	500173800EB00160	27/09/11 10:43 PM	900	1738.2	871.18	2609.38	13.6	6.8	20.4003	12.3	1	8.6	3.31	1.66	2.48
18	XIV-2812-7803760-IBM	1:FC_Port:7:1	500173800EB00170	27/09/11 10:43 PM	900	1739.63	874.19	2613.82	13.61	6.83	20.439	12.3	1	8.5	3.32	1.66	2.49
19	XIV-2812-7803760-IBM	1:FC_Port:8:1	500173800EB00180	27/09/11 10:43 PM	900	1737.87	873.19	2611.05	13.6	6.82	20.4277	12.3	1	8.5	3.32	1.66	2.49
20	XIV-2812-7803760-IBM	1:FC_Port:9:1	500173800EB00190	27/09/11 10:43 PM	900	1739.61	870.31	2609.93	13.61	6.81	20.4235	12.3	1.1	8.6	3.32	1.66	2.49
21	XIV-2812-7803760-IBM	1:FC_Port:4:1	500173800EB00140	27/09/11 10:58 PM	900	1752.1	879.29	2631.39	13.7	6.87	20.5752	12.3	1.1	8.6	3.34	1.67	2.5
22	XIV-2812-7803760-IBM	1:FC_Port:5:1	500173800EB00150	27/09/11 10:58 PM	900	1758.79	878.86	2637.65	13.72	6.87	20.5869	12.2	1	8.5	3.34	1.67	2.5
23	XIV-2812-7803760-IBM	1:FC_Port:6:1	500173800EB00160	27/09/11 10:58 PM	900	1753.36	879.14	2632.5	13.72	6.86	20.5854	12.3	1.1	8.6	3.34	1.67	2.5
24	XIV-2812-7803760-IBM	1:FC_Port:7:1	500173800EB00170	27/09/11 10:58 PM	900	1753.44	879.54	2632.99	13.71	6.86	20.5719	12.2	1	8.5	3.34	1.67	2.5
25	XIV-2812-7803760-IBM	1:FC_Port:8:1	500173800EB00180	27/09/11 10:58 PM	900	1754.36	875.27	2629.63	13.72	6.84	20.5572	12.2	1	8.5	3.34	1.67	2.5
26	XIV-2812-7803760-IBM	1:FC_Port:9:1	500173800EB00190	27/09/11 10:58 PM	900	1753.72	882.18	2635.9	13.71	6.88	20.5901	12.3	1	8.5	3.34	1.67	2.5
27	XIV-2812-7803760-IBM	1:FC_Port:4:1	500173800EB00140	27/09/11 11:13 PM	899	1737.81	869.12	2606.93	13.61	6.82	20.4265	12.4	1.1	8.6	3.32	1.66	2.49
28	XIV-2812-7803760-IBM	1:FC_Port:5:1	500173800EB00150	27/09/11 11:13 PM	899	1743.12	878.78	2621.9	13.61	6.87	20.4798	12.3	1.1	8.5	3.32	1.67	2.49
29	XIV-2812-7803760-IBM	1:FC_Port:6:1	500173800EB00160	27/09/11 11:13 PM	899	1739.3	870.05	2609.35	13.62	6.82	20.4453	12.4	1.1	8.6	3.32	1.66	2.49
30	XIV-2812-7803760-IBM	1:FC_Port:7:1	500173800EB00170	27/09/11 11:13 PM	899	1739.52	877.64	2617.16	13.62	6.85	20.4739	12.3	1	8.5	3.32	1.67	2.49
31	XIV-2812-7803760-IBM	1:FC_Port:8:1	500173800EB00180	27/09/11 11:13 PM	899	1738.06	875.22	2613.28	13.61	6.84	20.4429	12.3	1	8.5	3.32	1.66	2.49
32	XIV-2812-7803760-IBM	1:FC_Port:9:1	500173800EB00190	27/09/11 11:13 PM	899	1739.97	871.77	2611.74	13.62	6.84	20.4524	12.4	1	8.6	3.32	1.66	2.49
33	XIV-2812-7803760-IBM	1:FC_Port:4:1	500173800EB00140	27/09/11 11:28 PM	900	1737.46	870.49	2607.94	13.6	6.79	20.3911	12.3	1.1	8.6	3.32	1.65	2.48
34	XIV-2812-7803760-IBM	1:FC_Port:5:1	500173800EB00150	27/09/11 11:28 PM	900	1744.11	904.97	2649.08	13.63	6.96	20.595	12.2	1	8.4	3.32	1.69	2.5

Figure 6-30 Tivoli Storage Productivity Center data example

The data filters can be used to easily find maximum values. The spreadsheet can also be used to generate graphs. In our Tivoli Storage Productivity Center data example in Figure 6-30, the following graphs are frequently helpful:

- Total IOPS and total bandwidth.
- Read IOPS and read response times. If ports only, a read is a *Port send*.
- Write IOPS and write response times. If ports only, a write is a *Port receive*.

6.4 Performance evaluation

This section describes performance problem-solving techniques for XIV Storage System.

6.4.1 Problem-solving steps

The problem-solving steps consist of defining the problem definition, data colleciton, and performance analysis.

Problem definition

A good problem definition includes the following types of information:

- Which application is having performance problems? What is a description of the problem from the application perspective?
- What operating system performance characteristics indicate that the problem is XIV Storage System performance?
- What LUNs are dedicated to this application?

- How is the application server zoned?
- During what time period is the problem experienced?
- What is the XIV Storage System configuration?

There can always be more than one performance problem. It is best to define these performance issues separately but to note if they are related.

Data collection

The next thing to do is to collect good data. The data collected typically begins slightly before the problem begins and ends slightly after the performance problem duration. In this way, ideally, it is possible to see the abnormalities in the performance data (elevated response times, for example) begin and end. This task is fairly easy to do with the XIV Storage Management GUI performance statistics.

Typically, a performance issue is being experienced from a particular application server, so it is common to filter on one host or the specific volumes associated with that host.

Performance analysis

The first thing to consider when evaluating performance is the overall nature of the I/O characteristics. These characteristics typically include the IOPS, throughput measured in MBps, and latency measured in milliseconds (ms). Our sample XIV Storage System workload is shown in Figure 6-31.

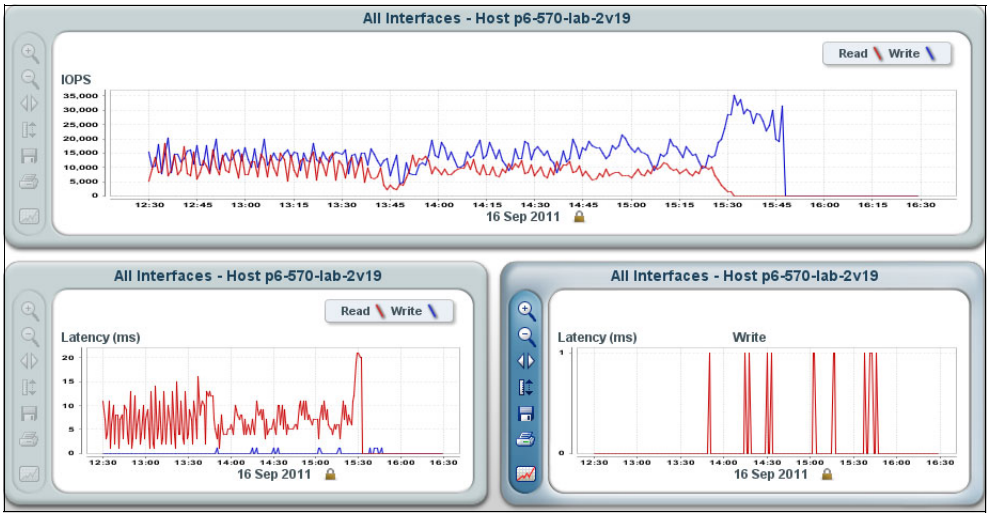


Figure 6-31 Host IOPS and response times

It is important during initial application testing and early production to document IOPS and response times when the application is running well. In this way, if an XIV Storage System performance issue is suspected, you can compare these known good performance characteristics to the new ones and see if there is a difference. If there is no difference, chances are that whatever is causing the problem, it is not the XIV Storage System.

Input/output operations per second and response time

When it comes to XIV Storage System performance analysis, there are a few questions:

- Is XIV providing the bandwidth or IOPS required by the application?
- Is XIV providing adequate read and write response times?

Because the XIV Storage System can have different configurations depending on the number of modules, each of these configurations has different performance capabilities. For a specified production workload, the different XIV Storage System configurations provide different performance characteristics. As the application performance demands increase, two things happen to the XIV Storage System performance characteristics:

1. The XIV Storage System provides more IOPS throughput (MBps) until it has reached its maximum performance capacity.
2. The XIV Storage System response time (latency) would tend to increase, especially as the XIV configuration approaches its maximum performance capacity.

Figure 6-32 is an example of a typical IOPS versus response time curve. It is based on theoretical I/O workload characteristics. Typical I/O characteristics include the following:

- IOPS
- I/O size for both reads and writes
- Cache hits for both reads and writes
- Percentage of reads versus writes
- Percentage of both reads and writes that are sequential

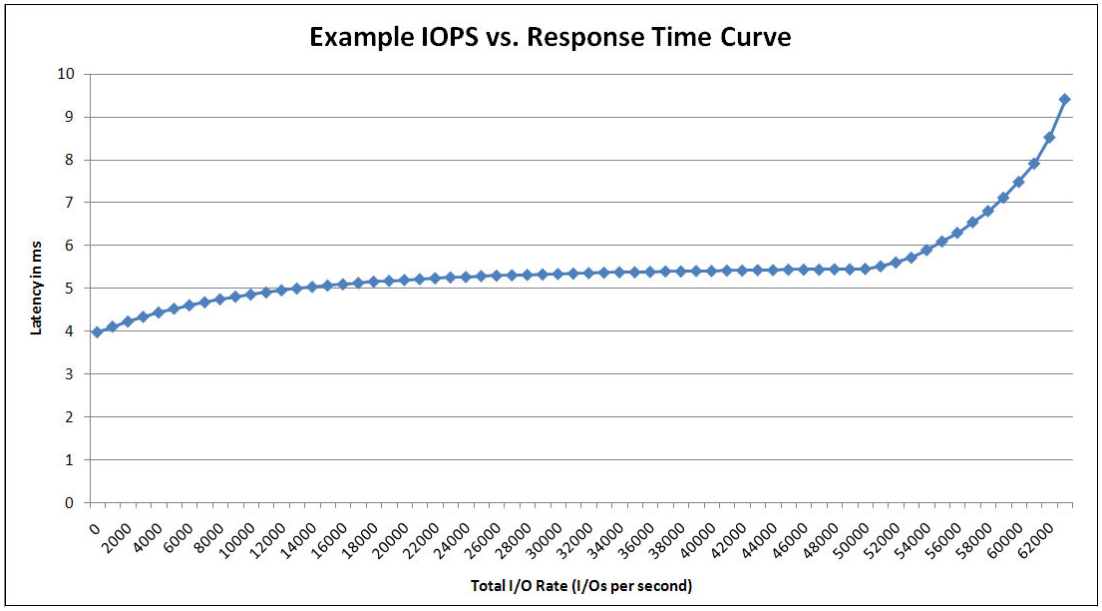


Figure 6-32 Example IOPS versus response time curve

The slope of the curve is almost flat as IOPS increase through most of the graph. But when the IOPS reaches 50,000, the latency begins to increase rapidly. This point is what is referred to as the *knee* of the IOPS versus response time curve. The shape of this curve is typical of all spinning disk systems. The actual IOPS versus latency numbers, the slope of the flat section, and the location of the knee would be different for each different disk system.

The main point that we want to make here is that the goal is to run production workloads in the horizontal section of the curve. As you monitor performance, if you begin to notice that response times are increasing significantly over their normal levels, it is time to determine the cause before the application owners begin to complain. There are many things that can cause response times to increase in this way, such as the following reasons:

- Fibre Channel or iSCSI SAN network issues
- Host performance issues causing the I/Os to be delivered slowly
- Pushing the XIV configuration to its performance limits (“operating in the knee”)

Prolonged periods of high response times associated with high IOPS or bandwidth use is typically an indication of a performance problem. This can occur because of failure to adhere to preferred practices at the logical configuration level at any or all stages of the I/O path and as a result of an insufficient storage hardware configuration to meet the peak I/O workload demands. It is critical to consider both avenues, and not just one in isolation. The former issue can be investigated by using the XIV Storage System GUI in conjunction with host configuration and monitoring tools to identify potential bottlenecks or non-ideal implementation practices based on the concepts presented in this chapter.

The other possible action is to evaluate the potential performance impact of upgrading and adding XIV Storage System hardware to meet the specific aggregate I/O workload requirements. This can be investigated, in part, by engaging IBM or an IBM Business Partner to perform a Disk Magic study. However, keep in mind that Disk Magic assumes that end-to-end logical configuration preferred practices have been implemented from the application layer to the storage subsystem layer. Therefore, attempting to use Disk Magic as the sole basis for diagnosing or isolating a performance issue might result in drawing an invalid conclusion and could prolong the duration of the problem.

Response times are a good measure of how well the XIV Storage System is performing. The XIV System GUI uses the term *latency*. They are both the same thing. The best way to determine whether response times are good or bad is to compare them to the values recorded when the application was known to be performing well. If the response times compare closely, it is likely that the XIV Storage System is still providing good performance. There are several things to consider when evaluating XIV Storage System response times:

- Do not be alarmed at sudden spikes in IOPS or response times. Most production applications are not sensitive enough to experience performance issues associated with spikes that last for only one reporting interval, which in the case of XIV Storage System is 1 minute.
- It is the prolonged elevated response times that most *transaction-based workloads* notice as poor performance. Prolonged can be 10 minutes or an hour or more because this measurement is dependent upon the application.
- Response times can be different depending on the type of workload the XIV Storage System is servicing. During *batch processing*, when throughput MBps is the primary performance objective, it is normal and common for batch response times to be higher for prolonged periods.
- Be careful about lumping response times into one number. Response times must be associated with either reads or writes. Notice the read and write response times in Figure 6-31 on page 321.

- Write response times for caching storage systems such as XIV Storage System are normally low because this time is the average time required for the write I/O operation to travel from the XIV Fibre Channel interface to cache. Write response times are highly dependent upon the size of the average write I/O operation. Very large write I/Os take longer for disk storage systems to process. For *transaction workloads* with write I/O sizes of 32 KB or less, it is common to observe write response times in the single-digit millisecond range. The best way to evaluate write response times is to compare them to the write response times recorded during good performance periods.
- Reads take longer than writes normally because some reads invariably must be retrieved from spinning disks and this task takes time. Caching storage systems mitigate this situation by detecting certain patterns in read I/O characteristics and prestaging data into cache. The XIV Storage System has exceptionally aggressive prestaging algorithms for sequential reads that place this data in DRAM cache.
- The XIV System also prestages small block random read data into the XIV Gen3 flash drives as described in 2.8.2, “Flash caching algorithm and concepts” on page 44.
- When read data is prestaged in flash or DRAM cache, the read request can be satisfied at cache I/O speeds. This is called a read hit. But invariably certain data, typically random in nature, must be retrieved from disk. This is called a read miss. Read response times are the weighted average of these fast read cache hits and the slow reads from disk.
- The amount of cache also has a significant impact on read cache hits. The more cache that you have, the more data you can put in there that can result in a read hit. XIV cache sizes can be described only as huge. A full rack has 720 GB of DRAM cache and 12 TB of flash cache.

  High read cache hit percentages are the most significant factor in good read performance and low read response times. The nature of the application read I/O is what defines the read cache hit percentage. If the XIV algorithms do a good job of prestaging data into flash or DRAM cache, the read cache hit percentage is high.

  XIV plays a large part in read cache hit percentages as well. XIV has exceptionally aggressive read data prestaging algorithms to get this read data into the huge XIV System cache. And XIV has a fixed number of spinning disks. But all these factors are fixed. When you see variations in the read hit percentage in the XIV statistics GUI, it is the application that is causing these variations.

  Read response times are highly dependent upon the size of the average read I/O operation. Very large read I/Os take longer for disk storage systems to process. For *transaction workloads* with read I/O sizes of 32 KB or less, it is common to observe read response times in the low double-digit millisecond range. Again, the best way to evaluate read response times is to compare them to the read response times recorded during good performance periods.

Interface module balance

Another important thing to check when evaluating XIV performance is the balance of I/O across the different interface modules. This evaluation requires an understanding of the SAN cabling and zoning. But if the preferred multipathing configuration described in 6.2.3, “Multipathing considerations” on page 297 is followed, every host uses every interface module equally, which can be seen in the XIV Storage Management GUI.

In Figure 6-33, interface modules 4 - 8 all have basically the same I/O characteristics. This balance is typical and desirable. interface module 9 has no I/O during most of this workload. Then, at 14:45, it starts to perform its share of the job of handling the I/O workload. In a normal situation, *all* interface modules work together *all* the time.

There are certain older operating systems, such as VMware ESX 3.5, that do not use more than one path to a volume. On these operating systems, ensure that you use more than one volume and that the paths for the volumes are balanced across all interface modules. You can accomplish this task by changing the preferred path for these volumes so that these operating systems cannot harm the performance of other operating systems or applications.

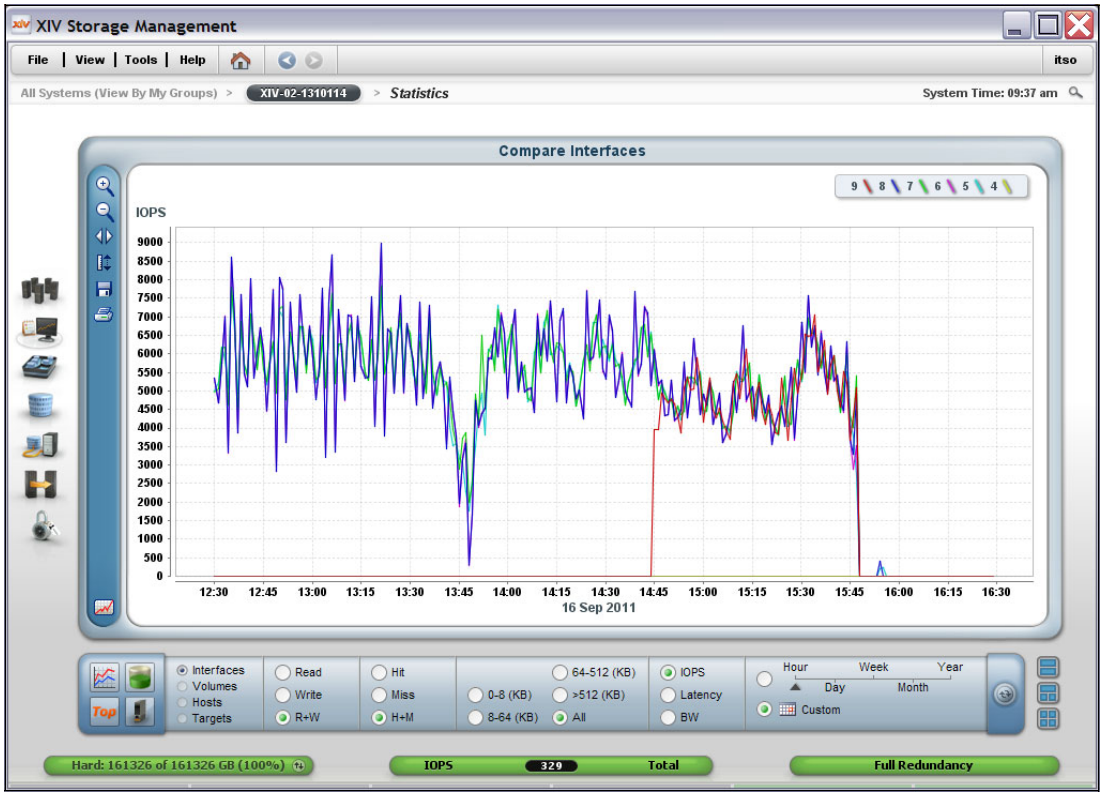


Figure 6-33 Interface module balance

7

Monitoring

In this chapter, we describe the various methods and functions that are available to monitor the IBM XIV Storage System. We also show how you can gather information from the system in real time, in addition to the self-monitoring, self-healing, and automatic alert functions implemented within the XIV Storage System Software.

This chapter covers the following topics:

- You can review or request the current system status and performance statistics at any time.
- You can set up alerts to be triggered when specific error conditions or problems arise in the system. Alerts can be conveyed as messages to a user, an email, and Simple Network Management Protocol (SNMP) trap or a Short Message Service (SMS) text to a mobile phone.
- Depending on the nature or severity of a problem, the system can automatically alert the IBM Remote Support Center, which immediately initiates the necessary actions to promptly repair the system. The Call Home feature must be configured and working properly for this feature to work.
- If you have purchased IBM Tivoli Storage Productivity Center, you can configure it to monitor and configure the XIV Storage System.
- You can use SNMP **`walk`** and **`get`** commands using the IBM supplied XIV SNMP Management Information Base (MIB) to monitor the health of the XIV Storage System.
- If you have IBM Systems Director, you can configure it to get SNMP events from the XIV Storage System.
- If you are using Microsoft System Center Operations Manager, you can use it to monitor your XIV Storage System.

The secure remote support feature allows remote monitoring and repair by IBM Suppot personnel. We describe the Call Home function, secure remote support, and repair procedures. Finally, the Host Attachment Kit now offers a data collection feature that can be used for problem analysis. It is described at the end of this chapter.

© Copyright IBM Corp. 2014. All rights reserved.

7.1 Monitoring with XIV Storage Management GUI and XCLI

The XIV Storage Management graphical user interface (GUI) and XIV Storage System Command-Line Interface (XCLI) include many features that allow you to monitor the system. These features include monitoring for alerts, events, and failed components.

7.1.1 Monitoring by using the Systems menu

You can show the health state, alerts, and events for multiple IBM XIV systems from the Systems menu. By selecting the Systems view, as shown in Figure 7-1, you can work with a particular storage system or show All Systems Alerts or All Systems Events for every XIV Storage System defined to the GUI. The Systems menu does not offer multiple selections when only one XIV Storage System is defined to the GUI.

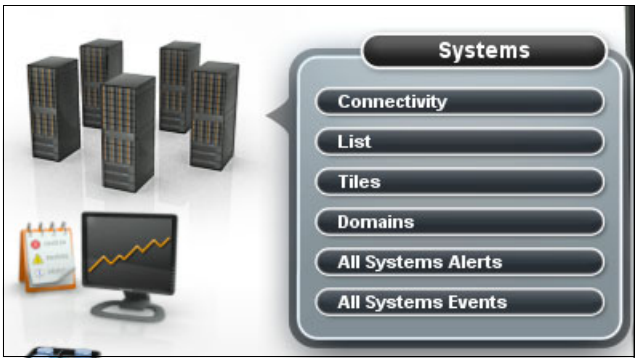


Figure 7-1 Systems menu view

The All Systems window

If you select **All Systems**, you are taken to a view showing all the defined IBM XIV systems. By default, the current input/output operations per second (IOPS) for each XIV Storage System is displayed. However, by right-clicking the green box at the top of any defined XIV Storage System, you can change all the IBM XIV systems to different metrics. This action is shown in Figure 7-2.

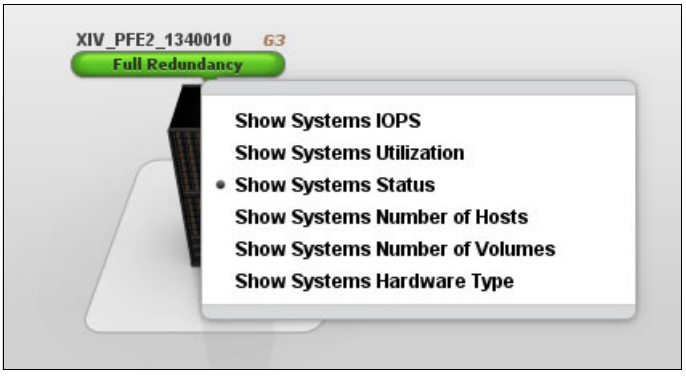


Figure 7-2 Changing the All Systems view

Possible metrics include IOPS, systems use (use of hard space as a percentage), systems status (the IBM XIV systems are fully redundant, redistributing, or rebuilding), number of defined hosts or number of defined volumes (with the number of snapshots shown in brackets), and system hardware type (XIV generation, disk type, and hardware configuration).

The All Alerts window

An alert is defined as a critical event or hardware failure in one of the monitored IBM XIV systems. If you open the All Alerts window, you see all the current alerts and failures for every system shown in the GUI. In Figure 7-3, you can see that there are several alerts for two different IBM XIV systems (although there are no current failures). If you double-click an alert, the next relevant window opens. For example, in Figure 7-3, there are several pool alerts. Therefore, double-clicking a pool alert changes the window to show the relevant pool highlighted on the pools view of the relevant XIV Storage System system.

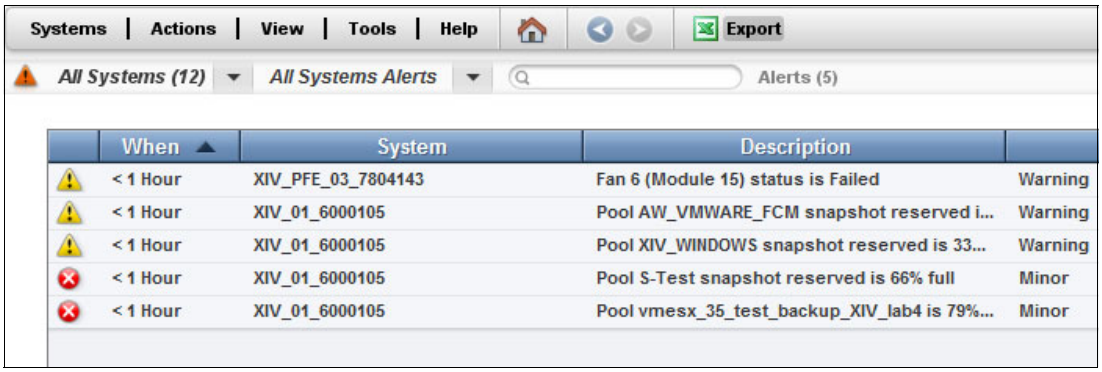


Figure 7-3 All Alerts window of the All Systems menu

All Systems Events window

If you select the **All Systems Events** view, a combined event log is displayed and shows all events for all systems that are defined to the GUI. In Figure 7-4, you can see events from three different systems in a combined view. This view can be sorted and filtered to reduce the number of events or change the order in which they are displayed.

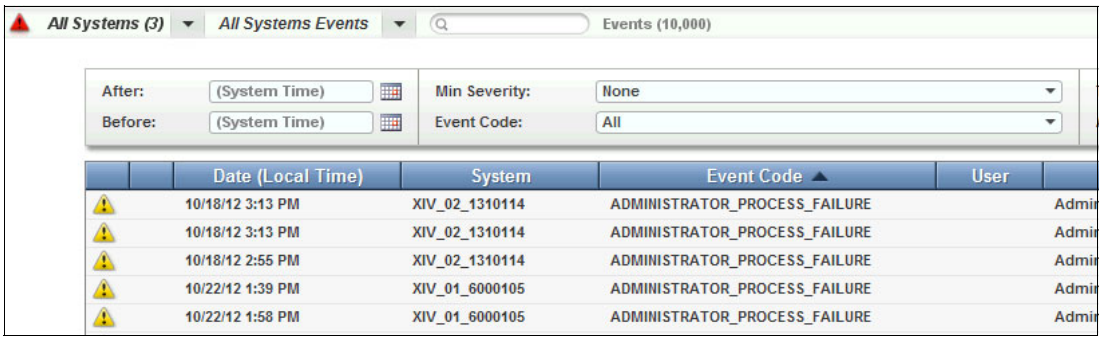


Figure 7-4 All Systems Events window

You can also access events and alerts for all IBM XIV systems from the View drop-down menu, as shown in Figure 7-5. You can view All Systems Alerts or All Systems Events or return to the All Systems view.

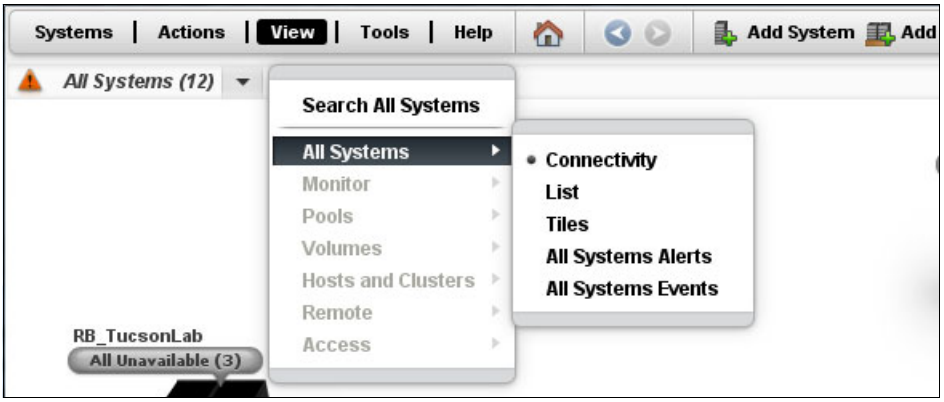


Figure 7-5 The View drop-down menu

7.1.2 Monitoring alerts for all defined IBM XIV systems with the GUI

The XIV Storage Management GUI allows users to work with alerts across multiple systems, regardless of which XIV Storage System is selected. In Figure 7-6, the cursor is hovering over the All Systems Alerts indicator in the lower right corner of the GUI. Three alerts across the three systems that are currently defined to the GUI are shown. You can access the alert information regardless of which system you are working with, or which window is displayed.

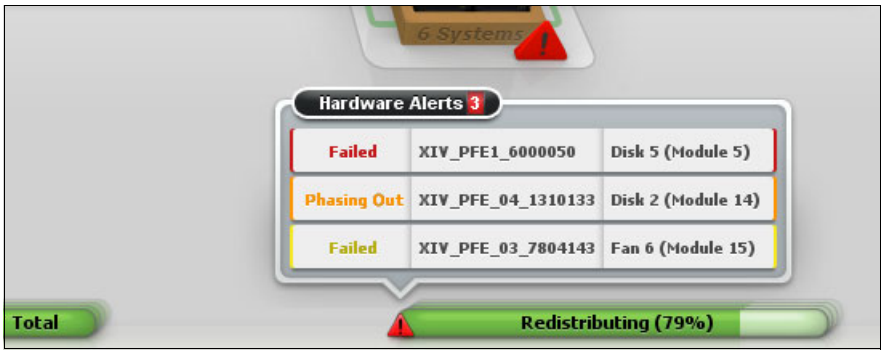


Figure 7-6 All Alerts window

In Figure 7-7, the cursor is hovering the alerts indicator for a single system, allowing the quantity of alerts to be displayed for just that system.

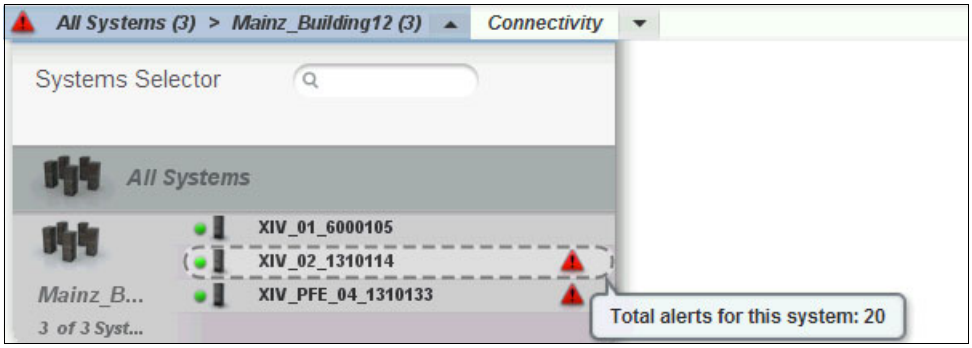


Figure 7-7 System level alert summary

As shown in Figure 7-8, selecting the alert icon for the system displays that system's specific alerts in a new window.

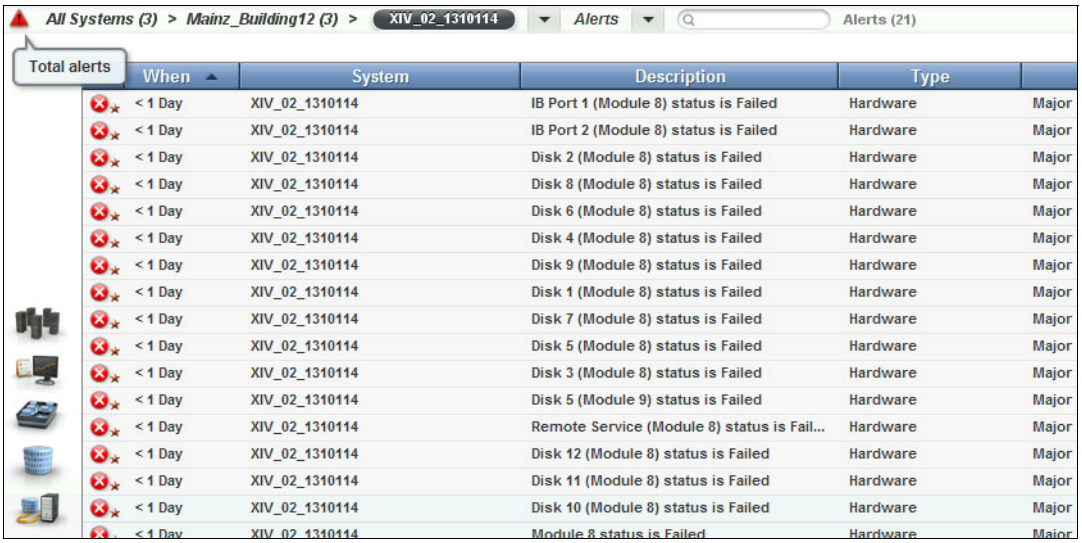

All Systems (3) > Mainz_Building12 (3) > XIV_02_1310114 Alerts Alerts (21)
Total alerts
When System Description Type
< 1 Day XIV_02_1310114 IB Port 1 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 IB Port 2 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 2 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 8 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 6 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 4 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 9 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 1 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 7 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 5 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 3 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 5 (Module 9) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Remote Service (Module 8) status is Fail... Hardware Major
< 1 Day XIV_02_1310114 Disk 12 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 11 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Disk 10 (Module 8) status is Failed Hardware Major
< 1 Day XIV_02_1310114 Module 8 status is Failed Hardware Major

Figure 7-8 Alerts window

7.1.3 Monitoring an individual XIV Storage System using the GUI

You can access the Alerts, Events, Statistics, and QoS Performance Class for the currently selected machine from the Monitor menu shown in Figure 7-9.

Figure 7-9 GUI monitor functions

Monitoring the system

Selecting **System** from the Monitor menu takes you to the System view, which is shown in Figure 7-10 on page 332 (this view is also the default or main GUI window for the selected system).

Figure 7-10 Monitoring the IBM XIV Storage System

The System view shows a graphical representation of the XIV Storage System rack with its components. You can click the curved arrow at the lower right of the picture of the rack to show a view of the patch panel. Figure 7-10 shows the patch panel for an XIV Storage System Gen3 Model 214.

You get a quick overview in real time about the system's overall condition and the status of its individual components. The display changes dynamically to provide details about a specific component when you hover the cursor over that component.

Status bar indicators at the bottom of the window, which are shown in Figure 7-11, indicate the overall operational levels of the XIV Storage System.

Figure 7-11 Status bar indicators

The following information describes the status bar indicators in Figure 7-11 on page 332:

- The first indicator on the left shows the amount of soft or hard storage capacity currently allocated to storage pools and provides alerts when certain capacity thresholds are reached. As the physical, or hard, capacity consumed by volumes within a storage pool passes certain thresholds, the color of this meter indicates that additional hard capacity might need to be added to one or more storage pools.

  Clicking the icons on the right side of the indicator bar that represent up and down arrows toggles the view between hard and soft capacity.

  Our example indicates that the system has a usable hard capacity of 158934 GB, of which 72% or 114296 GB is used.

  You can get more information and perform more accurate capacity monitoring by looking at storage pools (see 4.4.2, “Managing storage pools with the XIV Storage Management GUI” on page 142).
- The second indicator, in the middle, shows the number of IOPS.
- The third indicator on the far right shows the general system status and, for example, indicates when a redistribution is underway.

Monitoring events

To get to the Events window, select **Events** from the Monitor menu, as shown in Figure 7-12.

Extensive information and many events are logged by the XIV Storage System. The system captures entries about problems with various levels of severity, including warnings and other informational messages. These messages include information about logins, configuration changes, and the status of attached hosts and paths. All of the collected data can be reviewed in the Events window that is shown in Figure 7-12.

Because many events are logged, the number of entries is typically huge.

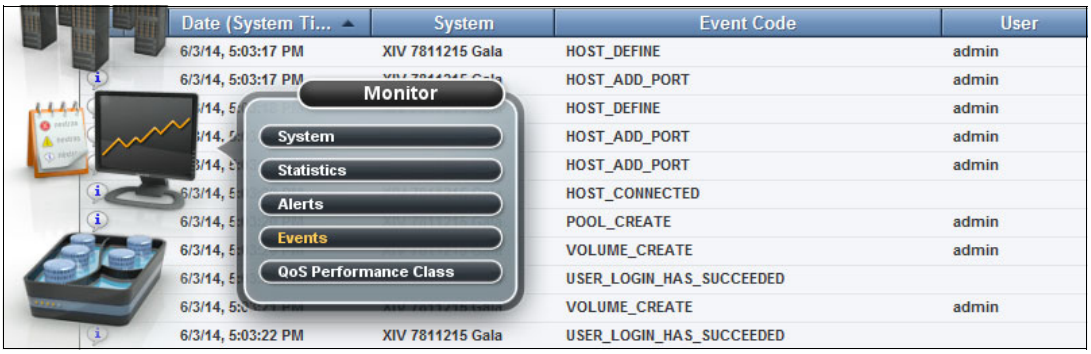


Figure 7-12 Events window

An option to filter the events logged can display a more useful and workable view. Without filtering the events, it might be difficult to find the entries for a specific incident or information. Figure 7-13 shows the possible filter options for the events.

Figure 7-13 Event filter

If you double-click a specific event in the list, you can get more information about that event, along with the best troubleshooting actions to take.

Figure 7-14 shows details for a major event where an internal network link is down. For this type of event, contact XIV Storage System Support immediately. Notification happens automatically if Call Home and remote support are enabled, as described in 7.5, “Call Home and remote support” on page 371. A problem record is generated with XIV Storage System Support, and an IBM Service Support Representative (SSR) contacts the client to report the problem and follow up with repair actions.

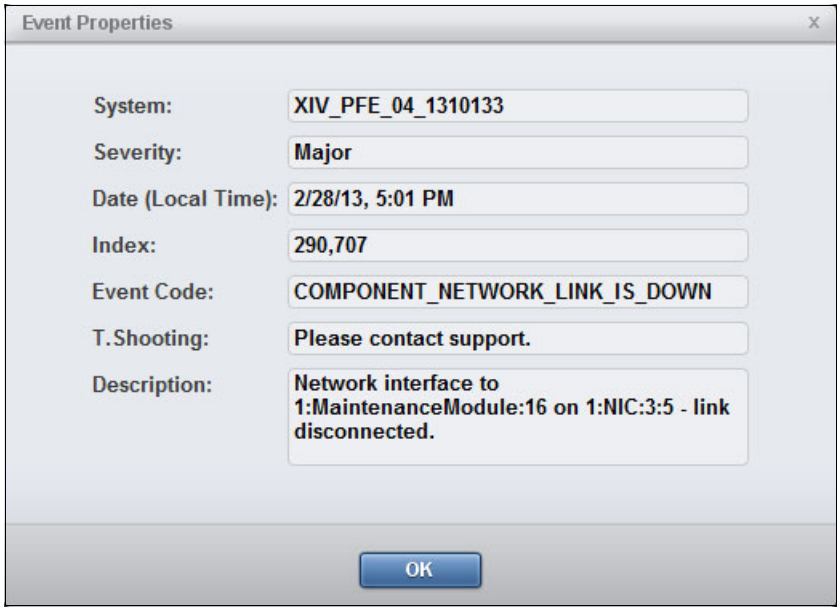


Figure 7-14 Event properties

Event severity

The events are classified into a level of severity depending on their effect on the system. Figure 7-15 gives an overview of the criteria and meaning of the various severity levels.

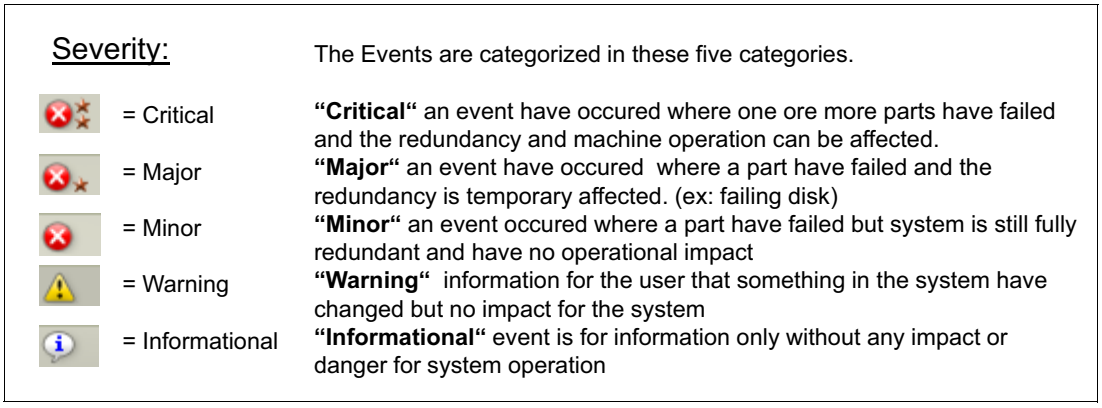

Severity:
The Events are categorized in these five categories.
= Critical
“Critical“ an event have occured where one ore more parts have failed and the redundancy and machine operation can be affected.
= Major
“Major“ an event have occured where a part have failed and the redundancy is temporary affected. (ex: failing disk)
= Minor
“Minor“ an event occured where a part have failed but system is still fully redundant and have no operational impact
= Warning
“Warning“ information for the user that something in the system have changed but no impact for the system
= Informational
“Informational“ event is for information only without any impact or danger for system operation

Figure 7-15 Event severity

Event notification configuration

The Events window offers a toolbar (Figure 7-16) that contains a setup wizard, with the ability to view and modify gateways, destinations, and rules. The wizard can also modify the email addresses for the XIV Storage System and export the events to a comma delimited (CSV) file for use in an external spreadsheet program.

Figure 7-16 Event rules configuration

Clicking the **Setup** icon starts the Events Configuration wizard, which guides you through the process to create gateways, add destinations, and define rules for event notification.

For more information about event notification rules, see “Setup notification and rules with the GUI” on page 359.

In addition to the standard system-generated events, a storage administrator can generate a **custom_event** with their own custom event description and wanted severity level using an XCLI command. This command is useful when creating scripts to automate a process. You can generate an informational event stating that a process has started or finished, as shown in Example 7-1.

Example 7-1 Using custom_event

```
>> custom_event description="User Initiated Backup Has Started" severity=INFORMATIONAL
command 0:
administrator:
    command:
        code = "SUCCESS"
        status = "0"
        status_str = "Command completed successfully"
aserver = "DELIVERY_SUCCESSFUL"

>> event_list code=CUSTOM_EVENT
Timestamp         Severity      Code         User          Description
------------------- --------------- -------------- ----------------- -----------------------------------
2010-10-13 14:15:48   Informational   CUSTOM_EVENT   itso              User Initiated Backup Has Started
```

Monitoring statistics

The statistics monitor provides information about the performance and workload of the XIV Storage System.

There is flexibility in how you can visualize the statistics. Options are selectable from a control pane at the bottom of the window, as shown in Figure 7-17.

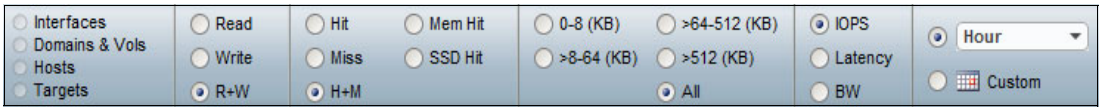


Figure 7-17 Filter pane for the statistics monitor

For information about performance monitoring, see Chapter 6, “Performance” on page 287.

QoS performance class

We previously introduced the quality of service (QoS) feature. QoS allows the XIV Storage System to deliver multiple service levels to hosts connected to the same system. See details in 4.7, “QoS feature” on page 170 on how to set QoS, by defining performance classes in terms of IOPS and bandwidth and then assigning specific hosts to a particular performance class. Besides host-based QoS the rate limiting can also be defined for pools and domains starting with XIV Storage Software v11.5.

Monitoring UPS status with the XIV Storage Management GUI

The XIV Storage Management GUI can be used to confirm if all three uninterruptible power supplies (UPSes) are currently operational. If you hover your cursor over a UPS, as shown in Figure 7-18, a window opens and shows the health status of that UPS and the Automatic Transfer Switch (ATS). The ATS generates system event messages if there are problems.

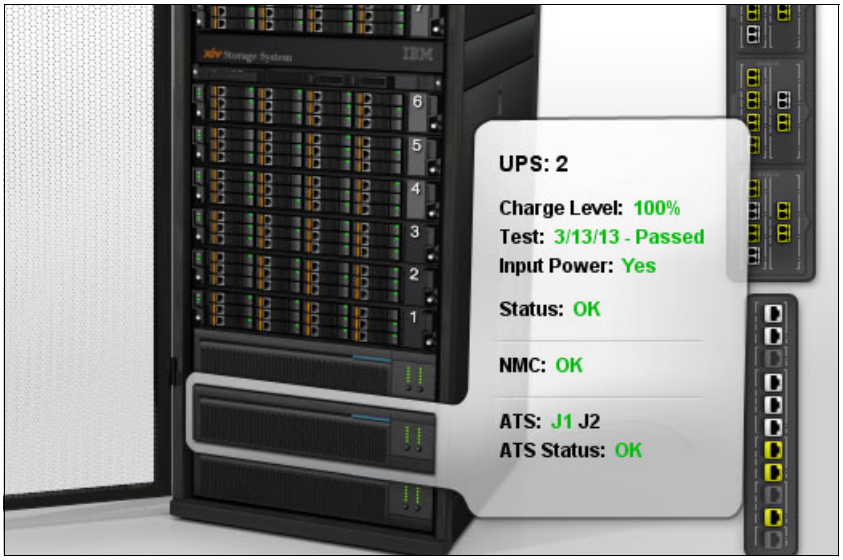


Figure 7-18 Monitoring UPS status

Monitoring InfiniBand switches

The XIV Storage Management GUI can be used to confirm the status of the InfiniBand switches. If you hover your cursor over a switch, as shown in Figure 7-19, a window opens and shows the status of that switch.

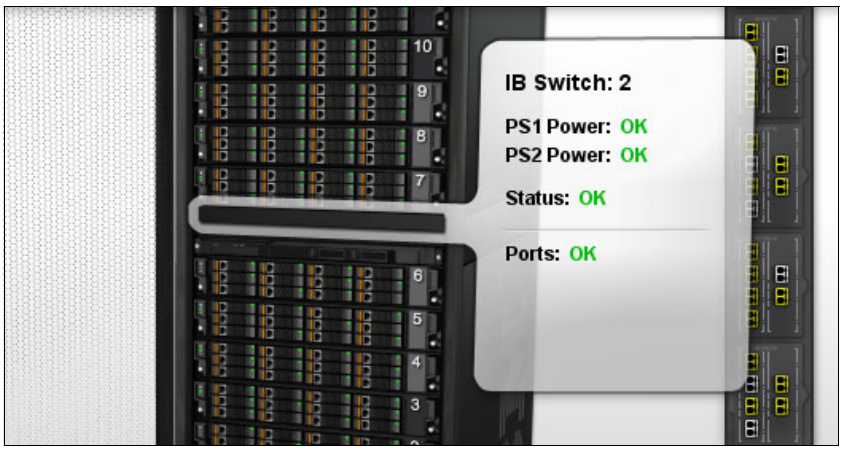


Figure 7-19 Monitoring UPS InfiniBand switches

Monitoring the Maintenance module

The XIV Storage Management GUI can be used to confirm the status of the Maintenance module. If you hover your cursor over a switch, as shown in Figure 7-20, a window opens and shows the status of that module.

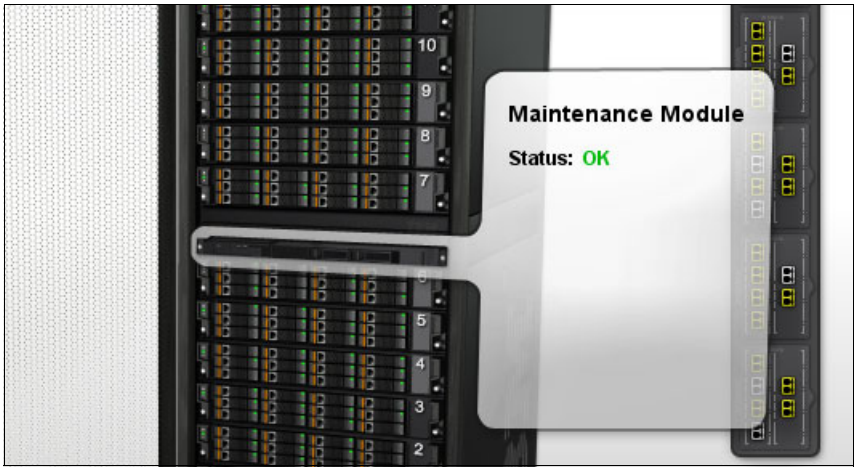


Figure 7-20 Monitoring the Maintenance module

Monitoring each module and disk status and temperature

The XIV Storage Management GUI can be used to monitor the status of each XIV Storage System module and each XIV Storage System disk. You can hover the cursor over a module in the location shown in Figure 7-21. The cursor changes to a magnifying glass icon. By hovering the cursor over this location, you can confirm the health of each module component and the module's temperature.

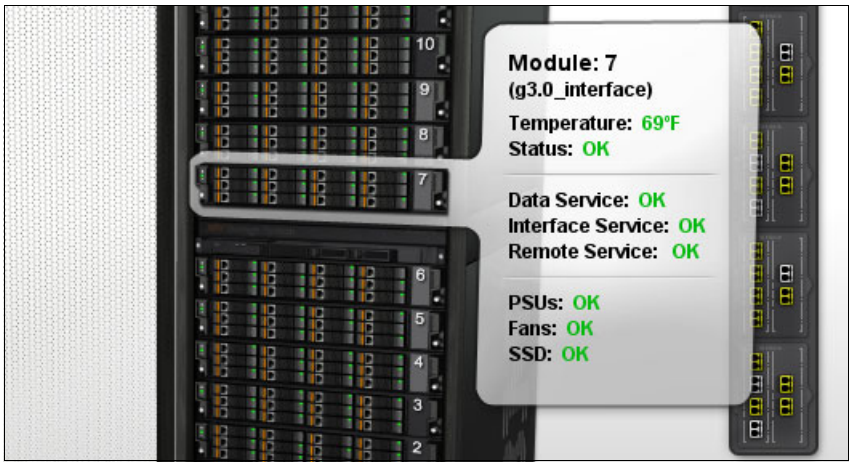


Figure 7-21 Module status

If you left-click the module, the module is removed from the rack and the status of the module components can be verified by hovering through the exploded drawing, as shown in Figure Figure 7-22.

Figure 7-22 Component status in module explosion drawing

Hover over a component to show the actual status of the single component in the module. The example in Figure 7-22 shows the status of PSU:1. The status and temperature of the module can be inspected also in this view. Click the thermometer to toggle between Celsius (C) and Fahrenheit (F).

If a module's temperature exceeds a default threshold, a small yellow or red triangle is displayed to indicate that the module temperature is higher than normal, as shown in Figure 7-23.

Figure 7-23 Module temperature threshold exceeded

If you hover your cursor over a disk, the status and temperature of the disk are shown (Figure 7-24).

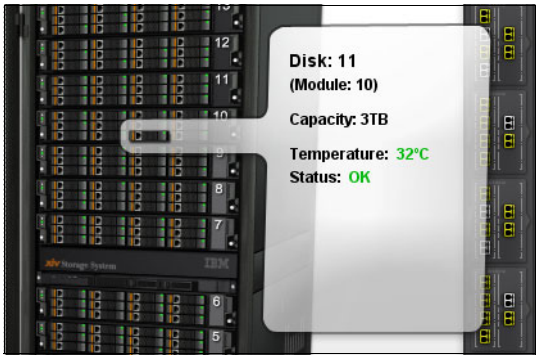


Figure 7-24 Disk status and temperature

Deferred disk replacement

XIV Storage System Gen3 has the concept of deferred disk replacement, which means that up to two failed disks can be marked for replacement later. If a third disk or another component fails, the XIV Storage System notifies IBM to replace all the failed components at the same time. A disk that has been placed into a deferred replacement state will not be red in color, but a shade of pink, as shown in Figure 7-25.

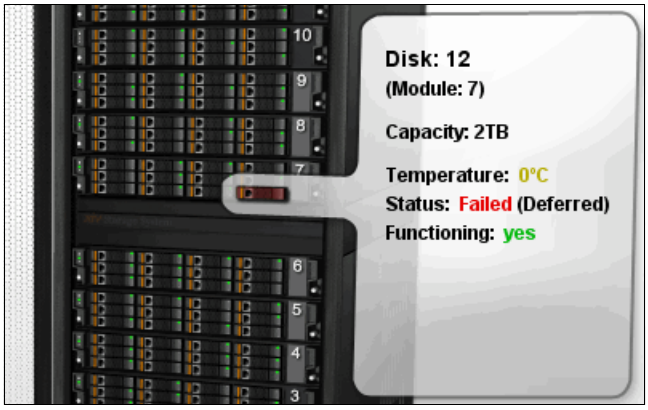


Figure 7-25 XIV Storage System Gen3 disk in a deferred replacement state

7.1.4 Monitoring with XIV Storage System Command-Line Interface

The XIV Storage System Command-Line Interface (XCLI) provides various commands to monitor the XIV Storage System and gather real-time system status, monitor events, and retrieve statistics. See 4.1, “XIV Storage Management software introduction” on page 112 for more information about how to set up and use the XCLI.

System monitoring

Various XCLI commands are available for system monitoring. For more information about these commands, see the *XCLI Utility User Manual* and *Commands Reference* books, which are available at the following website:

http://publib.boulder.ibm.com/infocenter/ibmxiv/r2/index.jsp

The **state_list** command, which is shown in Example 7-2, gives an overview of the general status of the system. In the example, the system is operational, data is fully redundant, no shutdown is pending, ssd_caching is enabled on the system level, and encryption is not supported.

Example 7-2 The state_list command

```
>>state_list
Category          Value
system_state      on
target_state      on
safe_mode         no
shutdown_reason   No Shutdown
off_type          off
redundancy_status Full Redundancy
ssd_caching       disabled
encryption        Not Supported
```

In Example 7-3, the **system_capacity_list** command shows an overview of used and free capacity, system-wide. In the example, both the hard and soft usable capacity is 161326 GB, with 42022 GB of free hard capacity and 39010 GB of free soft capacity. It also shows that all spare capacity is still available.

Example 7-3 The system_capacity_list command

```
>> system_capacity_list
Soft   Hard   Max_Pool_Size FreeHard FreeSoft SpareModules Spare Disks TargetSpareModules TargetSpareDisks
161326 161326 80654        42022    39010    1            3           1                  3
```

In Example 7-4, the **version_get** command shows the current version of the XIV Storage System code installed on the system. Knowing the current version of your software assists you in determining when upgrades are required.

Example 7-4 The version_get command

```
>>version_get
Version
11.2.0.a
```

In Example 7-5, the **time_list** command is used to retrieve the current time from the XIV Storage System. This time is normally set at the time of installation. Knowing the current system time is required when reading statistics or events. In certain cases, the system time might differ from the current time (at the user's location). Therefore, knowing when something occurred according to the system time assists with debugging issues.

Example 7-5 The time_list command

```
>> time_list
Time       Date         Time Zone   Daylight Saving Time
23:05:28   2013-03-19   GMT         no Time       Date         Time Zone
```

System components status

In this section, we present various XCLI commands that are used to obtain the status of specific system components, such as disks, modules, or adapters.

The **component_list** command, which is shown in Example 7-6, gives the status of all hardware components in the system. The filter option **filter=<FAILED | NOTOK>** is used to return information only about failing components. The first example shows a failed disk in module 4 slot 9, and the second example shows that all components are in an OK status.

Example 7-6 The component_list command

```
>> component_list filter=NOTOK
Component ID  Status  Currently Functioning
1:Disk:4:9    Failed  no

>> component_list filter=NOTOK
No components match the given criteria
```

As shown in Example 7-7 on page 341, the **disk_list** command provides more in-depth information for any individual disk in the XIV Storage System, which might be helpful in determining the root cause of a disk failure. If the command is issued without the disk parameter, all the disks in the system are displayed.

Example 7-7 The disk_list command

```
>> disk_list disk=1:Disk:13:10
Component ID   Status   Currently Functioning   Capacity (GB) Vendor   Model          Size      Serial
1:Disk:13:10   OK       yes                     2TB           IBM-XIV  ST32000444     1878632   9WM4PHXG
```

Example 7-8 shows a disk that is failed but in a deferred replacement state. For more details about this state, see “Deferred disk replacement” on page 339.

Example 7-8 Disk in deferred replacement

```
>> component_list filter=NOTOK
Component ID   Status   Currently Functioning
1:Disk:7:12    Failed   yes
```

In Example 7-9, the **module_list** command shows details about the modules themselves. If the **module** parameter is not provided, all the modules are displayed. In addition to the status of the module, the output describes the number of disks, number of Fibre Channel (FC) ports, and number of IP network Small Computer System Interface (iSCSI) ports.

Example 7-9 The module_list command

```
>> module_list module=1:Module:4
Component ID   Status   Currently Functioning   Target Status   Type             Data Disks   FC Ports
1:Module:4     OK       yes                                     g3.0_interface   12           4
iSCSI Ports   Temperature
2             22
```

It is possible to use the **-x** XCLI parameter with commands to provide even more information about a component. The **-x** parameter shows the output of a command in XML format. You can see the additional information as displayed in Example 7-10.

Example 7-10 Using -x parameter to show additional information in XML format

```
>> module_list module=1:Module:4 -x
<XCLIRETURN STATUS="SUCCESS" COMMAND_LINE="module_list module=1:Module:4 -x">
  <OUTPUT>
            <module id="5760a900003">
                <component_id value="1:Module:4"/>
                <status value="OK"/>
                <currently_functioning value="yes"/>
                <requires_service value=""/>
                <service_reason value=""/>
                <target_status value=""/>
                <type value="g3.0_interface"/>
                <disk_bay_count value="12"/>
                <fc_port_count value="4"/>
                <ethernet_port_count value="4"/>
                <io_allowed value="yes"/>
                <io_enabling_priority value="0"/>
                <serial value="MBC0944037G3E64"/>
                <original_serial value="MBC0944037G3E64"/>
                <part_number value="98Y3335"/>
                <original_part_number value="98Y3335"/>
                <usm_version value="3.6RC1"/>
                <bmc_version value="5.06.00000000"/>
                <bios_version value="5.00.00000000"/>
```

```
            <fpga_version value="1.00.00000000"/>
            <ses_version value="3.1.0.18"/>
            <pdb_firmware value="4.01"/>
            <pcm_1_firmware value="1.06"/>
            <pcm_2_firmware value="1.06"/>
            <fan_controller_firmware value="UCD90910|2.0.9.0000|100520 (2.2.0)"/>
            <battery_firmware value="0.00"/>
            <sas_version value="0a030002h"/>
            <infiniband_hca_version value="2.9.1810"/>
            <cna_version value="2.9.1200"/>
            <memory_gb value="24"/>
            <temperature value="24"/>
            <normalized_temperature value="22"/>
            <chassis_serial value="SHM0984831PUJDT"/>
            <chassis_part_number value="0984831-01"/>
            <electronics_serial value="FTS0941313G167X"/>
            <electronics_part_number value="0941313-04"/>
          </module>

  </OUTPUT>
</XCLIRETURN>
```

In Example 7-11, the **ups_list** command describes the status of the uninterruptible power supply (UPS) component. It provides details about when the last test was performed and the results. Equally important is the current battery charge level. A battery that is not fully charged can be a cause of problems in case of power failure.

The output of the **ups_list** command is broken into two lines for easier reading.

Example 7-11 The ups_list command

```
>> ups_list
>>ups_list
Component ID   Status   Currently Functioning   Input Power On   Runtime Remaining   Battery Charge Level
1:UPS:1        OK       yes                     yes              13                  100
1:UPS:2        OK       yes                     yes              21                  100
1:UPS:3        OK       yes                     yes              20                  100
-
Last Self Test Date   Last Self Test Result   Monitoring Enabled   UPS Status
03/13/2013            Passed                  yes                  ON_LINE
03/13/2013            Passed                  yes                  ON_LINE
03/14/2013            Passed                  yes                  ON_LINE
```

Example 7-12 shows the **switch_list** command that is used to show the status of the switches. An XIV Storage System second-generation Model A14 has Ethernet switches. An XIV Storage System Gen3 Model 114 reports an InfiniBand switch.

Example 7-12 The switch_list command

```
>> switch_list
1:IB_Switch:1   OK       OK                 0            2013/03/20   00:09:48   OK
1:IB_Switch:2   OK       OK                 0            2013/03/20   00:09:56   OK
-
Power Supply #2   Problematic Fans   Problematic Temperatures   Problematic Voltages
OK                0                  0                          0
OK                0                  0                          0
```

The **psu_list** command that is shown in Example 7-13 lists all the power supplies in each of the modules. You can show an individual power supply unit (PSU) by using the **psu=** parameter. With this command, you can also see the actual power consumption of the PSU (model 214 is equipped with high efficiency power supply units).

Example 7-13 The psu_list command

```
>> psu_list
Component ID   Status   Currently Functioning   Hardware Status   Power
1:PSU:1:1      OK       yes                     OK                143.12W
1:PSU:1:2      OK       yes                     OK                104.14W
.
.
1:PSU:15:1     OK       yes                     OK                143.12W
1:PSU:15:2     OK       yes                     OK                100.20W
```

By using the **-x parameter**, more information about the actual values can be displayed. See Example 7-14.

Example 7-14 The psu_list -x psu=1:PSU:1:1 command

```
>> psu_list -x psu=1:PSU:1:1
<XCLIRETURN STATUS="SUCCESS" COMMAND_LINE="psu_list -x psu=1:PSU:1:1">
  <OUTPUT>
            <psu id="3e4031000l4">
                <component_id value="1:PSU:1:1"/>
                <status value="OK"/>
                <currently_functioning value="yes"/>
                <requires_service value=""/>
                <service_reason value=""/>
                <hardware_status value="OK"/>
                <serial value="PMA0963415G01GR"/>
                <part_number value="0963415-05"/>
                <original_serial value="PMA0963415G01GR"/>
                <original_part_number value="0963415-05"/>
                <fru_descriptor value="DC"/>
                <original_fru_descriptor value="DC"/>
                <firmware_version value="1.06"/>
                <voltage_12v value=" 12.18V"/>
                <voltage_5v value="  5.11V"/>
                <current value=" 12.07A"/>
                <power value="147.01W"/>
            </psu>

  </OUTPUT>
</XCLIRETURN>
```

Events

Events can also be managed with XCLI commands. Various commands are available to list, filter, close, and send notifications for the events. There are many commands and parameters available.

We illustrate just a few of the options of the **event_list** command. Various parameters can be used to sort and filter the output of the **event_list** command. See Table 7-1 on page 344 for a list of the most commonly used parameters.

Table 7-1 The event_list command parameters

Name	Description	Syntax and example
max_events	Lists a specific number of events.	<event_list max_events=100>
after	Lists events after the specified date and time.	<event_list after=2008-08-11 04:04:27>
before	Lists events before the specified date and time.	<event_list before 2008-08-11 14:43:47>
min_severity	Lists events with the specified and higher severities.	<event_list min_severity=major>
alerting	Lists events for which an alert was sent or for which no alert was sent.	<event_list alerting=no> <event_list alerting=yes>
cleared	Lists events for which an alert was cleared or for which the alert was not cleared.	<event_list cleared=yes> <event_list cleared=no>

These parameters can be combined for better filtering. In Example 7-15, two filters were combined to limit the amount of information displayed. The first parameter, **max_events**, allows only five events to be displayed. The second parameter is the date and time that the events must not exceed. In this case, the event occurred approximately 1.5 minutes before the cutoff time.

Example 7-15 The event_list command with two filters combined

```
>> event_list max_events=5 before=2013-03-08.18:00:00
Timestamp             Severity        Code                            User   Description
2013-03-08 16:23:09   Informational   USER_LOGIN_HAS_SUCCEEDED                User 'xiv_maintenance' from IP
2013-03-08 16:23:41   Informational   TARGET_CONNECTION_ESTABLISHED           Target named 'XIV_04_1340008'
2013-03-08 16:23:41   Informational   TARGET_CONNECTION_ESTABLISHED           Target named 'XIV_04_1340008'
2013-03-08 16:24:14   Warning         USER_HAS_FAILED_TO_RUN_COMMAND          User 'xiv_administrator' from
2013-03-08 16:28:35   Informational   USER_LOGIN_HAS_SUCCEEDED                User 'xiv_development' from IP
```

The event list can also be filtered for severity. Example 7-16 shows all the events in the system that contain a severity level of Major and all higher levels, such as Critical.

Example 7-16 The event_list command filtered for severity

```
>> event_list min_severity=Major max_events=5
Timestamp             Severity   Code                                 User   Description
2013-03-18 13:05:20   Major      TARGET_DISCONNECTED                         Target named 'XIV_04_1340008' i
2013-03-18 13:05:43   Major      TARGET_LINK_DOWN_BEYOND_THRESHOLD           Target named 'XIV_04_1340008'
2013-03-19 14:27:22   Major      TARGET_DISCONNECTED                         Target named 'XIV_04_1340008'
2013-03-19 14:27:40   Major      TARGET_LINK_DOWN_BEYOND_THRESHOLD           Target named 'XIV_04_1340008'
```

Certain events generate an alert message and do not stop until the event has been cleared. These events are called *alerting events* and can be viewed by the XIV Storage Management GUI or XCLI using a separate command. After the alerting event is cleared, it is removed from this list, but it is still visible with the **event_list** command. See Example 7-17.

Example 7-17 The event_list_uncleared command

```
>> event_list_uncleared
No alerting events exist in the system
```

Monitoring statistics

The statistics gathering mechanism is a powerful tool. The XIV Storage System continually gathers performance metrics and stores them internally. Using the XCLI, data can be retrieved and filtered by using many metrics. Example 7-18 provides an example of gathering the statistics for 10 days, with each interval covering an entire day. The system is given a time stamp as the ending point for the data. Because of the magnitude of the data being provided, it is best to redirect the output to a file for further post-processing. For a more in-depth view of performance, see Chapter 6, “Performance” on page 287.

Example 7-18 Statistics for 10 days

```
>> statistics_get count=10 interval=1 resolution_unit=day end=2013-03-19.14:00:00
```

The **usage_get** command is a useful tool to provide details about the current use of pools and volumes. The system saves the usage every hour for later retrieval. This command works the same as the **statistics_get** command. You specify the time stamp to begin and end the collection and the number of entries to collect. In addition, you need to specify the pool name or the volume name. See Example 7-19.

Example 7-19 The usage_get command by pool

```
>> usage_get pool=ITSO_3 max=5 start=2013-09-19.08:00:00 end=2013-09-19.19:00:00

Time                 Volume Usage (MiB)   Snapshot Usage (MiB)
2013-09-19 08:00:00  0                    0
2013-09-19 09:00:00  0                    0
2013-09-19 10:00:00  1920087              0
2013-09-19 11:00:00  1920087              0
2013-09-19 12:00:00  1920087              0
```

The use is displayed in MB. Example 7-20 shows that the volume is using 1920 GB of space.

Example 7-20 The usage_get command by volume

```
>> usage_get vol=ITSO_Blade5_Perf max=5 start=2013-09-19.08:00:00 end=2013-09-19.19:00:00

Time                 Volume Usage (MiB)   Snapshot Usage (MiB)
2013-09-19 08:00:00  0                    0
2013-09-19 09:00:00  0                    0
2013-09-19 10:00:00  1920077              0
2013-09-19 11:00:00  1920082              0
2013-09-19 12:00:00  1920082              0
```

7.1.5 XIV audit event logging

The XIV Storage System uses a centralized event log. For any command that has been executed that leads to a change in the system, an event entry is generated and recorded in the event log. The object creation time and the user are also logged as object attributes.

The event log is implemented as a circular log and is able to hold a set number of entries. When the log is full, the system wraps back to the beginning. If you need to save the log entries beyond what the system normally holds, you can issue the **event_list** XCLI command and save the output to a file.

Event entries can be viewed by the XIV Storage Management GUI, XCLI commands, or by notification. A flexible system of filters and rules allows you to generate customized reports and notifications. For details about how to create customized rules, see 7.1.9, “Defining notification rules” on page 350.

7.1.6 Viewing events in the XIV Storage Management GUI

The XIV Storage Management GUI provides a convenient and easy to use view of the event log. To get to the view shown in Figure 7-26, right-click the Monitor icon from the main XIV Storage Management GUI window and click **Events** from the menu.

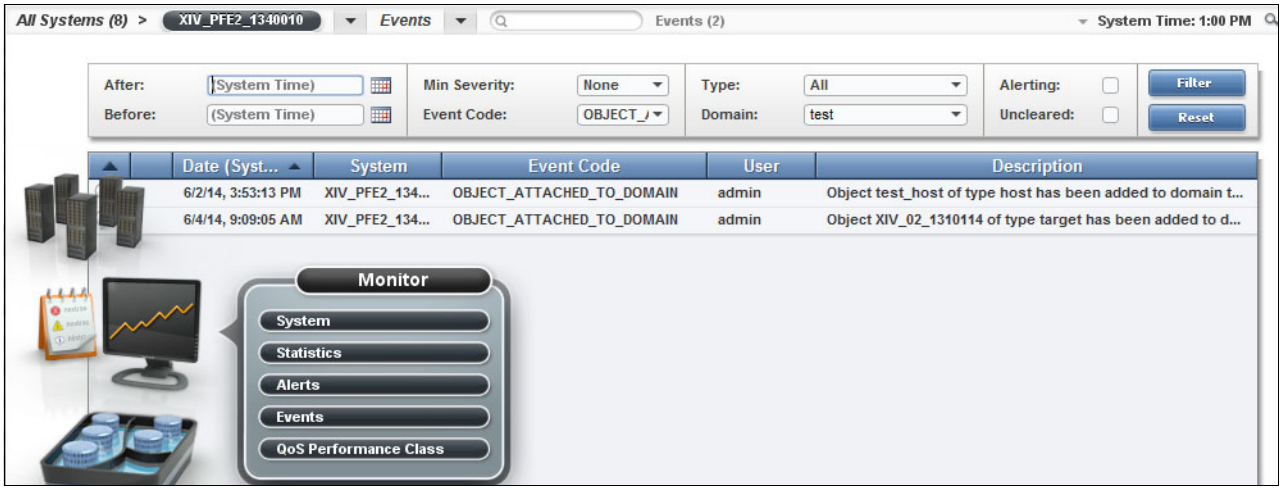


Figure 7-26 XIV Storage Management GUI Events main view

The window is split into two sections:

- The top part contains the management tools, such as wizards, in the menu bar, and a series of input fields and drop-down menus that act as selection filters.
- The bottom part is a table displaying the events according to the selection criteria. Use the table tile bar or headings to enable or change sort direction.

The system progressively loads the events into the table. A progress indicator is visible at the lower right of the table, as shown in Figure 7-27.

Figure 7-27 Loading events into the table

7.1.7 Event attributes

This section describes all the available event types, event codes, and their severity levels.

Severity levels

You can select one of six possible severity levels as the minimal level to be displayed:

- None: Includes all severity levels
- Informational: Changes, such as volume deletion, size changes, or host multipathing
- Warning: Volume usage limits reach 80%, failing message sent
- Minor: Power supply power input loss, volume usage over 90%, and component TEST failed
- Major: Component failed (disk), user system shutdown, volume and pool usage 100%, UPS on battery, or Simple Mail Transfer Protocol (SMTP) gateway unreachable
- Critical: Module failed or UPS failed

Event codes

For a list of event codes, see the *IBM XIV Storage System User Manual*, GC27-3914.

Event types

The following event types can be used as filters (specified with the **`object_type`** parameter in the XCLI command):

cons_group	Consistency group
destgroup	Event destination group
dest	Event notification group
dm	Data migration
domain	Domain
host	Host
map	Volume mapping
mirror	Mirroring
pool	Pool
rule	Rule
smsgw	SMS gateway
smtpgw	SMTP gateway
target	Fibre Channel or iSCSI connection
volume	Volume mapping
cluster	Cluster
ip_interface	IP interface
ldap_conf	LDAP configuration
meta_data_object	Metadata events
sync_schedule	Schedules
user	User
user_group	User group
ldap_server	LDAP server
modules_status	Modules status
xmirror	3-way mirror

Domain event

Client can select the appropriate domain name, if any defined, from the Domain pull-down menu and can get events list related to that corresponding domain as shown in Figure 7-26 on page 346

7.1.8 Viewing events using the XCLI

Table 7-2 on page 348 provides a list of all the event-related commands available in the XCLI. This list covers setting up notifications and viewing the events in the system. For a more in-depth description of system monitoring, see Chapter 7, “Monitoring” on page 327.

Table 7-2 XCLI: All event commands

Command	Description
custom_event	Generates a custom event.
dest_define	Defines a new destination for event notifications.
dest_delete	Deletes an event notification destination.
dest_list	Lists event notification destinations.
dest_rename	Renames an event notification destination.
dest_test	Sends a test message to an event notification destination.
dest_update	Updates a destination.
destgroup_add_dest	Adds an event notification destination to a destination group.
destgroup_create	Creates an event notification destination group.
destgroup_delete	Deletes an event notification destination group.
destgroup_list	Lists destination groups.
destgroup_remove_dest	Removes an event notification destination from a destination group.
destgroup_rename	Renames an event notification destination group.
event_clear	Clears alerting events.
event_list	Lists system events.
event_list_uncleared	Lists uncleared alerting events.
event_redefine_threshold	Redefines the threshold of a parameterized event.
smsgw_define	Defines a Short Message Service (SMS) gateway.
smsgw_delete	Deletes an SMS gateway.
smsgw_list	Lists SMS gateways.
smsgw_prioritize	Sets the priorities of the SMS gateways for sending SMS messages.
smsgw_rename	Renames an SMS gateway.
smsgw_update	Updates an SMS gateway.
smtpgw_define	Defines an SMTP gateway.
smtpgw_delete	Deletes a specified SMTP gateway.
smtpgw_list	Lists SMTP gateways.
smtpgw_prioritize	Sets the priority of which SMTP gateway to use to send emails.
smtpgw_rename	Renames an SMTP gateway.
smtpgw_update	Updates the configuration of an SMTP gateway.
rule_activate	Activates an event notification rule.
rule_create	Creates an event notification rule.
rule_deactivate	Deactivates an event notification rule.
rule_delete	Deletes an event notification rule.

Command	Description
rule_list	Lists event notification rules.
rule_rename	Renames an event notification rule.
rule_update	Updates an event notification rule.
domain_list	Lists all domains.
domain_list_users	Lists users in the system per domain.
domain_list_objects	Lists objects in the system per domain

Event_list command and parameters

The following syntax is for the **event_list** command:

```
event_list
[ max_events=MaxEventsToList ] [ after=TimeStamp ] [ before=TimeStamp ]
[ min_severity=<INFORMATIONAL|WARNING|MINOR|MAJOR|CRITICAL> ]
[ alerting=<yes|no|all> ] [ cleared=<yes|no|all> ] [ code=EventCode ] [
object_type=<cons_group|destgroup|dest|dm|host|map|mirror|pool|rule|smsgw|smtpgw|t
arget|volume|cluster|ip_interface|ldap_conf|meta_data_object|sync_schedule|user|us
er_group|ldap_server|modules_status|elicense|ipsec_connection|ipsec_tunnel|cross_c
ons_group,...> ] [ internal=<yes|no|all> ] [ beg=BeginIndex ] [ end=EndIndex ]
[ count_all=<yes|no> ]
```

XIV Storage System Command-Line Interface examples

To illustrate how the commands operate, the **event_list** command displays the events currently in the system. Example 7-21 shows the first few events logged in our system.

Example 7-21 XCLI viewing events

```
C:\XIV>xcli -c "XIV 6000050" event_list
Timestamp             Severity        Code          User    Description
2011-10-04 17:00:38   Informational   USER_DEFINED  admin   A user with name
'itso_testuser1' and category applicationadmin was defined.
2011-10-04 17:01:05   Informational   POOL_RENAME   admin   Storage Pool with name 'ITSO'
was renamed 'ITSO2'.
2011-10-04 17:51:00   Informational   MAP_VOLUME    admin   Volume with name
'itso_app01_vol01' was mapped to LUN '1' for host with name 'itso_app02'.
2011-10-04 17:51:58   Informational   MAP_VOLUME    admin   Volume with name
'itso_app01_vol02' was mapped to LUN '2' for host with name 'itso_app02'.
.......
```

Example 7-22 shows the command for listing all instances when the user was updated. The USER_UPDATED event is generated when a user's password, email, or phone number is modified. In this example, the **-t** option is used to display specific fields, such as index, code, description of the event, time stamp, and user name. The description field provides the ID that was modified, and the user field is the ID of the user performing the action.

Example 7-22 View USER_UPDATED event with the XCLI

```
C:\XIV>xcli -c "XIV 6000050" -t index,code,description,timestamp,user_name event_list
code=USER_UPDATED
Index Code           Description                              Timestamp        User
1089  USER_UPDATED User with name 'admin' was updated.       2009-07-10 21:45:50 admin
1091  USER_UPDATED User with name 'admin' was updated.       2009-07-10 21:45:51 admin
1092  USER_UPDATED User with name 'technician' was updated. 2009-07-10 21:45:52
xiv_development
1093  USER_UPDATED User with name 'technician' was updated.  2009-07-10 21:45:53
xiv_development
```

7.1.9 Defining notification rules

Example 7-23 describes how to set up a rule in the XCLI to notify the storage administrator when a user's access control has changed. The rule has four event codes that generate a notification. The events are separated with commas with no spaces around the commas. If any of these four events are logged, the XIV Storage System uses the "relay" destination to issue the notification.

Example 7-23 Setting up an access notification rule using the XCLI

```
C:\XIV>xcli -c "XIV 6000050" rule_create rule=test_rule
codes=ACCESS_OF_USER_GROUP_TO_CLUSTER_REMOVED,ACCESS_OF_USER_GROUP_TO_HOST_REMOVED
,ACCESS_TO_CLUSTER_GRANTED_TO_USER_GROUP,ACCESS_TO_HOST_GRANTED_TO_USER_GROUP
dests=email_dest
Command executed successfully.
```

A simpler example is setting up a rule notification for when a user account is modified. Example 7-24 creates a rule on the XIV Storage System called *ESP* that sends a notification whenever a user account is modified on the system. The notification is transmitted through the relay destination.

Example 7-24 Create a rule for notification with the XCLI

```
C:\XIV>xcli -c -c "XIV 6000050" rule_create rule=user_update codes=USER_UPDATED
dests=email_dest
Command executed successfully.
```

The same rule can be created in the XIV Storage Management GUI. For more details about configuring the system to provide notifications and setting up rules, see Chapter 7, "Monitoring" on page 327.

7.2 Monitoring using the IBM XIV Mobile Dashboard

IBM has released a Mobile Dashboard application that can be used to securely monitor the performance and health of your XIV Storage System. This application is available for the Apple iPad (either iPad 1, iPad 2, or the newer iPad).

7.2.1 Installing the IBM XIV Mobile Dashboard

You need an Apple iOS device (either iPad or iPhone) and a valid Apple ID to download the IBM XIV Mobile Dashboard application from the App Store. It is a no-charge application. It can be found by searching the iTunes store, as shown in Figure 7-28, or by going to the following website:

https://itunes.apple.com/us/app/ibm-xiv-mobile-dashboard-universal/id503500546?mt=8

Figure 7-28 IBM XIV Mobile Dashboard at the Apple Store

Using the IBM XIV Mobile Dashboard with an Apple iPad

When you start the IBM XIV Mobile Dashboard, you have the choice to either run in Demo Mode or to connect to an actual XIV Storage System. Demo Mode can be accessed by selecting the Demo Mode option indicated by an arrow in the lower-right corner of the display, as shown in Figure 7-29 on page 352.

Figure 7-29 IBM XIV Mobile Dashboard Demo Mode

To log on to an actual XIV Storage System, use a valid user, password, and IP address, as shown in Figure 7-30.

Figure 7-30 Logging on to an XIV Storage System from the IBM XIV Mobile Dashboard

After you are connected to the XIV Storage System, you can view volume performance or host performance.

Each window also shows the health and redundancy state of the XIV Storage System. The iOS device has a screen rotation capability; therefore, the output can be shown in either landscape or portrait mode.

The IBM XIV Mobile Dashboard behaves differently depending on which mode you use to show the output.

If you are having issues with screen rotation, consult the following website:

http://support.apple.com/kb/HT4085

Volume performance

Volume performance is shown when the iPad is in landscape mode, as shown in Figure 7-31. From this window, you can see up to 120 seconds worth of performance for the highlighted volume. Use your finger to rotate the arrow on the blue volume icon to switch the display between IOPS, bandwidth (in megabytes per second (MBps)), and latency (in milliseconds (ms)). The data redundancy state of the XIV Storage System is shown in the upper-right corner (in this example, it is in Full Redundancy, but it can be Rebuilding or Redistributing). Swipe to the left to go to the Hosts window.

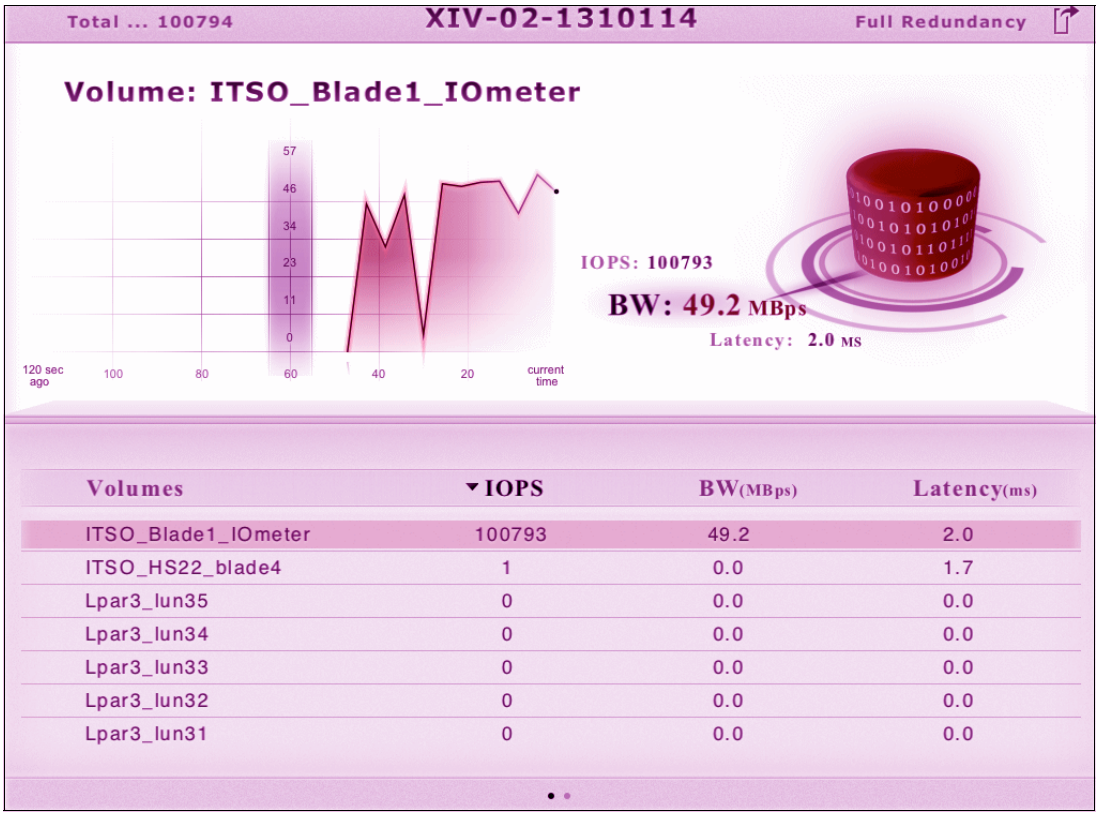


Figure 7-31 IBM XIV Mobile Dashboard Volume view

If the Mobile Dashboard is displayed in portrait mode (by rotating your iOS device), a list of up to 27 volumes is displayed instead.

Host view

The Host performance window is displayed when the iOS device is in landscape mode, as shown in Figure 7-32. From this window, you can see up to 120 seconds worth of performance for the highlighted host. Use your finger to rotate the arrow on the purple host icon to switch the display between IOPS, bandwidth (in MBps), and latency (in ms). The data redundancy state of the XIV Storage System is shown in the upper-right corner (in this example, it is in Full Redundancy, but it can potentially also be Rebuilding or Redistributing). Swipe to the right to go to the Volumes window.

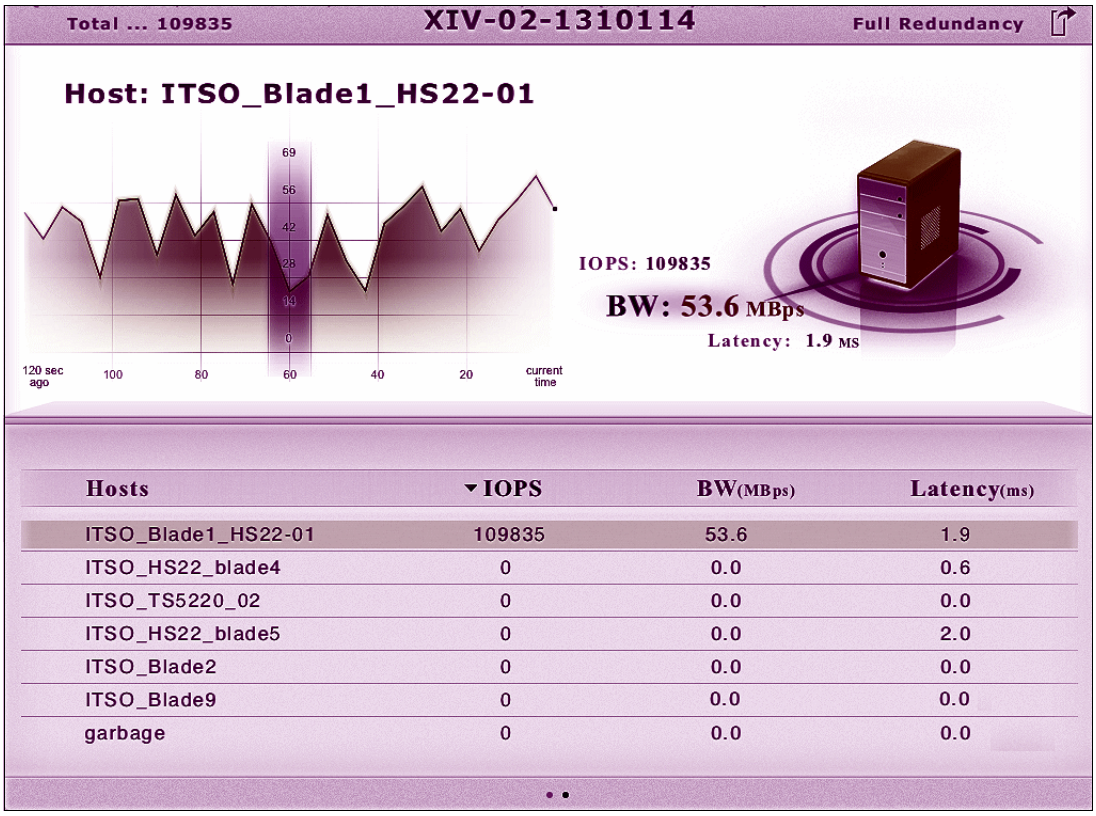


Figure 7-32 IBM XIV Mobile Dashboard Host view

If the IBM XIV Mobile Dashboard is operated in portrait mode (by rotating your iPad), a list of up to 27 hosts is displayed instead.

From either the Volume view or the Host window, you can log off from the IBM XIV Mobile Dashboard by using the icon in the upper-right corner of the display. When you restart the application, it recalls the IP address and user that were last used, but not the password. The password must be entered again.

Mobile Dashboard for the Apple iPhone

The IBM XIV System Mobile Dashboard is available for the Apple iPhone. . The demo mode allows you to try the available monitoring features without actually connecting to an IBM XIV System. The login sequence is the same as detailed before.

There is a performance monitoring summary window for monitoring overall IBM XIV System performance (see Figure 7-33) and the ability to view specific performance measurements.

Figure 7-33 iPhone login and system summary

You can monitor IOPS, bandwidth, and latency by host and by volume (see Figure 7-34).

Figure 7-34 iPhone monitoring screens

7.2.2 Mobile Dashboard for Android devices

The IBM XIV Mobile Dashboard is now available for Android. It has functions that are similar to the Apple iOS devices versions. The Demo Mode allows you to try the available monitoring features without actually connecting to an IBM XIV System. You can download the Mobile Dashboard app from Google Playstore:

https://play.google.com/store/apps/details?id=com.ibm.xiv.mobile.android&hl=en

At the welcome screen, either enter Demo Mode or add an XIV system. In Figure 7-35, the login screen and Demo Mode are shown.

Figure 7-35 XIV mobile dashboard for Android login and Demo Mode

There are performance monitoring windows for overall IBM XIV System performance, host performance, and volume performance. You can also view specific performance measurements. In Figure 7-36 on page 357, the system performance and volume performance are depicted.

Figure 7-36 XIV mobile dashboard for Android System and Volumes performance

7.2.3 IBM Storage Mobile dashboard

IBM has released an IBM Storage Mobile Dashboard application that can be used to securely monitor the performance and health of your IBM Storage System, it supports the following IBM storage systems:

- IBM Storwize® V7000/V7000 Unified
- IBM Flex System® V7000 Storage Node
- IBM Storwize V5000
- IBM Storwize V3700/V3500
- IBM SAN Volume Controller (SVC)
- IBM FlashSystem™
- IBM XIV Storage System
- IBM TS4500 Tape Library

You need an Apple iOS device (either iPad or iPhone) and a valid Apple ID to download the IBM Storage Mobile Dashboard application from the App Store. It is a no-charge application.

It can be found in the iTunes store by going to the following website: (see Figure 7-37)

https://itunes.apple.com/se/app/ibm-storage-mobile-dashboard/id677826483?mt=8

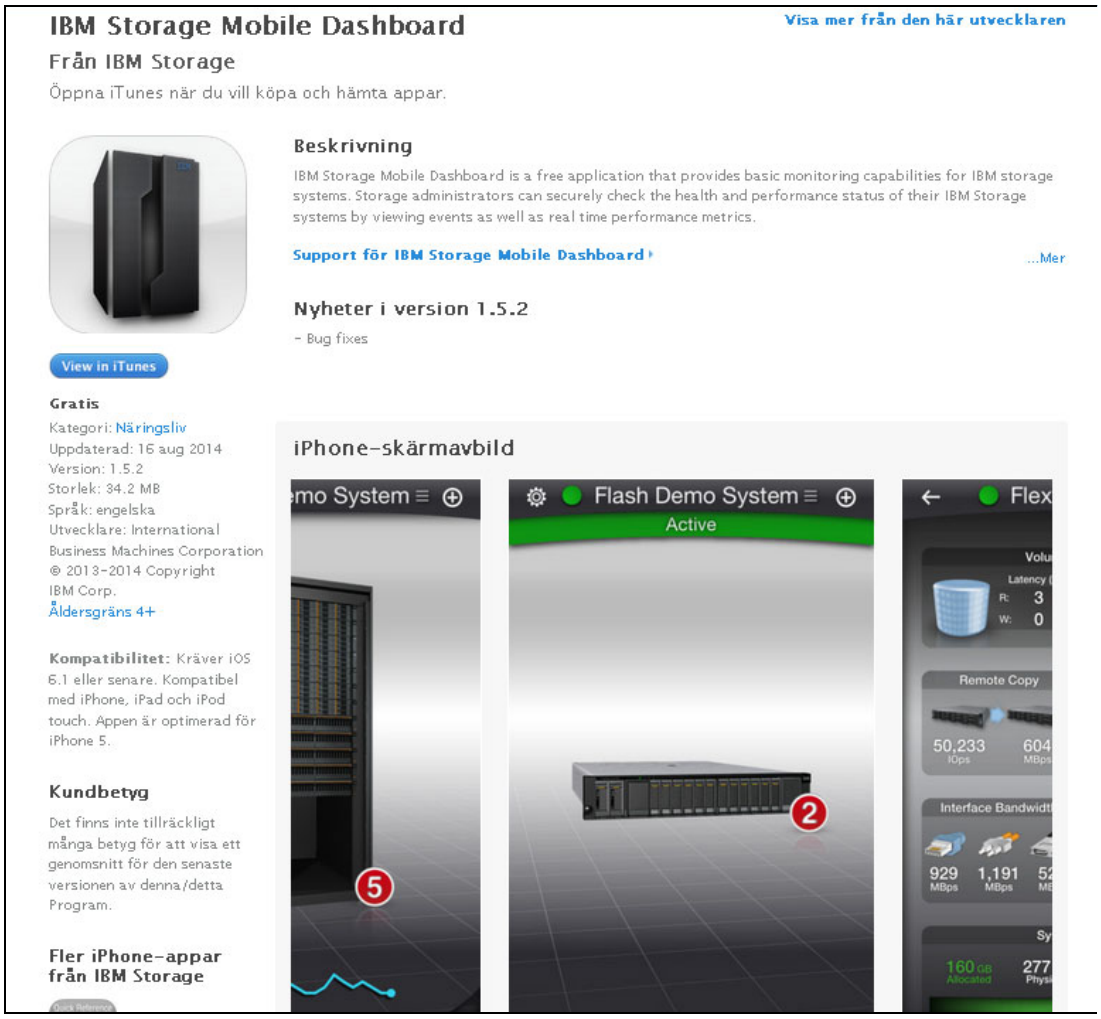

IBM Storage Mobile Dashboard
Visa mer från den här utvecklaren
Från IBM Storage
Öppna iTunes när du vill köpa och hämta appar.
Beskrivning
IBM Storage Mobile Dashboard is a free application that provides basic monitoring capabilities for IBM storage systems. Storage administrators can securely check the health and performance status of their IBM Storage systems by viewing events as well as real time performance metrics.
Support för IBM Storage Mobile Dashboard
...Mer
Nyheter i version 1.5.2
- Bug fixes
View in iTunes
Gratis
Kategori: Näringsliv
Uppdaterad: 16 aug 2014
Version: 1.5.2
Storlek: 34.2 MB
Språk: engelska
Utvecklare: International Business Machines Corporation
© 2013-2014 Copyright IBM Corp.
Åldersgräns 4+
Kompatibilitet: Kräver iOS 6.1 eller senare. Kompatibel med iPhone, iPad och iPod touch. Appen är optimerad för iPhone 5.
Kundbetyg
Det finns inte tillräckligt många betyg för att visa ett genomsnitt för den senaste versionen av denna/detta Program.
Fler iPhone-appar från IBM Storage
iPhone-skärmavbild
Flash Demo System
Active
Remote Copy
50,233
929
1,191

Figure 7-37 IBM Storage Mobile Dashboard

7.3 XIV Storage System event notification

The XIV Storage System allows you to send alerts by email, SMS messages, and SNMP traps.

7.3.1 Configuring the XIV Storage System to send notifications

You can configure the XIV Storage System using flexible rules to ensure that notification is sent to the correct person, or group of people, according to the various parameters. This event notification is similar to, but not quite the same, as XIV Storage System Call Home, which is described in 7.5, “Call Home and remote support” on page 371.

7.3.2 Setting up event notification

Configuration options are available from the XIV Storage Management GUI. You have the flexibility to create a detailed events notification plan based on specific rules. This flexibility allows the storage administrator to decide, for example, where to direct alerts for various event types. All these settings can also be done with XCLI commands.

Setup notification and rules with the GUI

To set up email or SMS notification and rules, complete the following steps:

1. From the XIV Storage Management GUI main window, select the **Monitor** icon. From the Monitor menu, select **Events** to open the Events window, as shown in Figure 7-38.

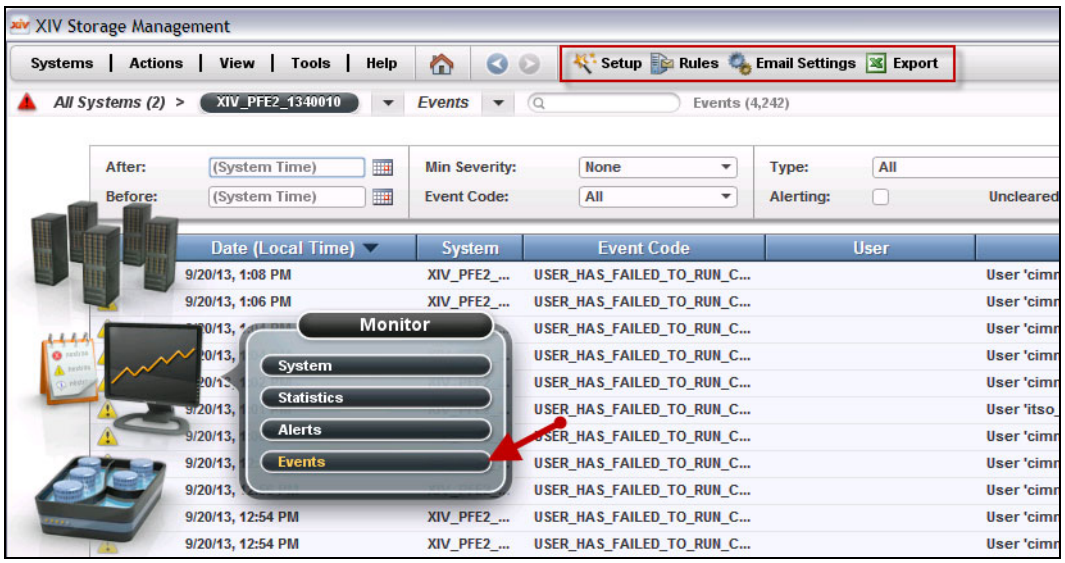


Figure 7-38 Setup notification and rules

2. From the toolbar, click **Setup** to start the Events Configuration wizard. The wizard guides you through the configuration of gateways, destinations, and rules. From the initial Welcome panel, click **Next** or **Gateway** to open the Events Configuration - Gateway window, as shown in Figure 7-39.

Figure 7-39 Define Gateway window

3. Click **Define Gateway**. The Gateway Create Welcome window that is shown in Figure 7-40 opens. Click **Next**. The Gateway Create - Select gateway type window opens, as shown in Figure 7-40.

Figure 7-40 Select gateway type

4. When the wizard prompts for the type of the gateway, either SMTP for email notification or SMS if an alert or information will initiate an SMS, click either **SMTP** or **SMS**.

Gateways: An SMTP gateway must be defined before you can define an SMS gateway because SMS messages are sent from the XIV Storage System in the form of an email.

The next steps differ for SMTP and SMS. Our illustration from this point forward is for SMTP. However, the steps to go through for SMS are self-explanatory and are described in “Setting up notifications and rules using the XCLI” on page 366.

To proceed with SMTP, click **Next**.

5. Enter the gateway name of the SMTP gateway and click **Next**. Enter the IP address or DNS name of the SMTP gateway for the gateway address and click **Next**. The SMTP Sender Email Address window opens, as shown in Figure 7-41 on page 361.

Figure 7-41 SMTP Sender Email Address window

6. Set the sender email address. You can use the default, or enter a new address. If there are email problems, such as the wrong email address, a response email is sent to this address. Depending on how your email server is configured, you might need to use an authorized address to ensure correct delivery of notifications. Click **Finish**.

   The Create the Gateway summary window opens, as shown in Figure 7-42.

Figure 7-42 Create the Gateway: Summary

7. Review the information you entered on this window. If all is correct, click **Create**. If not, click **Back** until you are at the information that needs to be changed. Or, select the buttons on the left to take you directly to the information that needs to be changed.

   Next, the Events Configuration wizard guides you through the setup of the destinations, where you configure email addresses or SMS receivers. Figure 7-43 on page 362 shows the Destination window of the Events Configuration wizard.

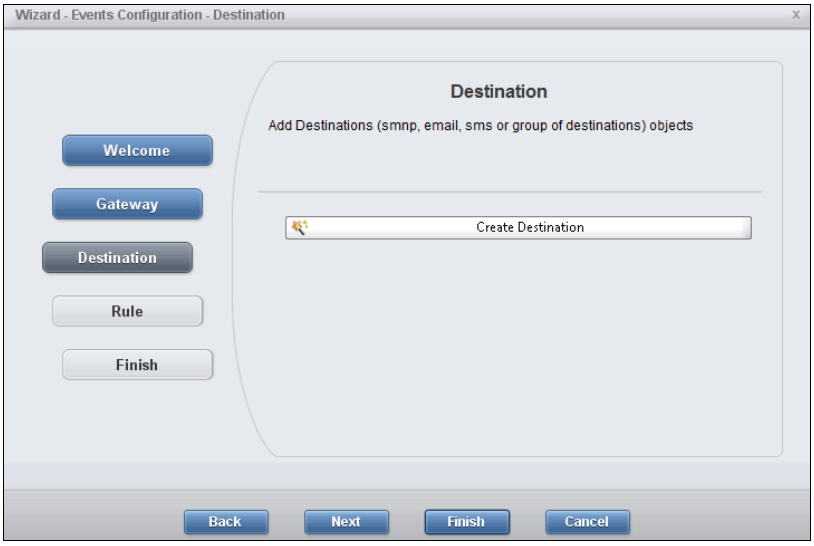


Figure 7-43 Add Destination window

8. Click **Create Destination** to open the Welcome window. Then, click **Next** to proceed. The Select Destination type window opens, as shown in Figure 7-44. On this window, you configure the following settings:

   – Type: Event notification destination type can be either a destination group (containing other destinations), SNMP manager for sending SNMP traps, email address for sending email notification, or mobile phone number for SMS notification:

   - SNMP
   - EMAIL
   - SMS
   - Group of Destinations

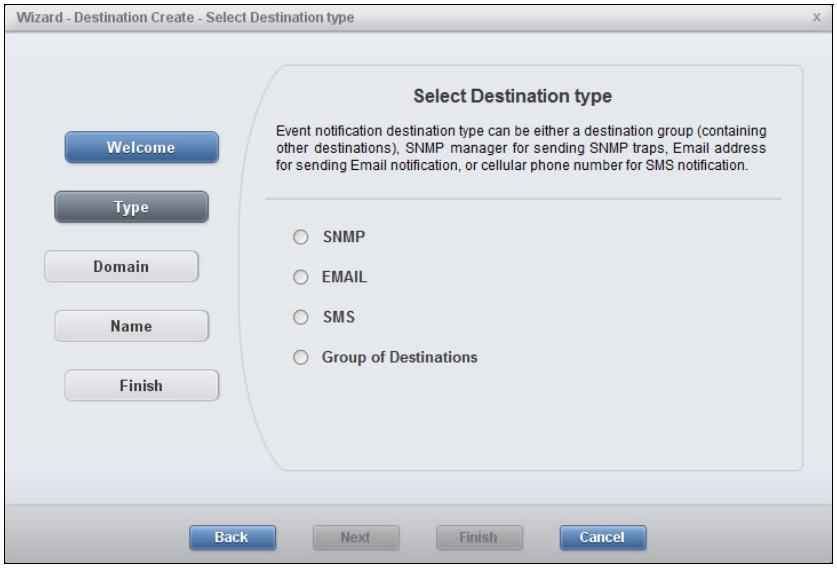


Figure 7-44 Select Destination type window

– Depending on the selected type, the remaining configuration information that is required differs, but is self-explanatory.

The final step in the Events Creation wizard is creating a rule. A rule determines what notification is sent. It is based on event severity, event code, or both.

9. Click **Create Rule**, as shown in Figure 7-45.

Figure 7-45 Create Rule window

10. On the Welcome window, click **Next**. The Rule Create - Rule name window opens, as shown in Figure 7-46.

Figure 7-46 Rule name window

11. To define a rule, configure the following settings:
    - Rule Domain:Select the new rule's domain. Select **All** to associate the rule with all of your domains. The default name is "no domain," which means that the corresponding rule will not be attached to specified domain.
    - Rule Name: Enter a name for the new rule. Names are case-sensitive and can contain letters, digits, or the underscore character (_). You cannot use the name of an already defined rule.
    - Rule condition setting: Select **Severity only** if you want the rule to be triggered by severity. Select **Event Code only** if you want the rule to be triggered by event. Or, select **Both Severity and event code** for events that might have multiple severities depending on a threshold of certain parameters.
    - Select the severity trigger: Select the minimum severity to trigger the rule's activation. Events of this severity or higher trigger the defined rule.
    - Select the event code trigger: Select the event code to trigger the rule's activation from the drop-down list. The list of event codes is predefined and cannot be modified. See the *XCLI User Manual*, GC27-2213, for a list of all the available event codes and descriptions to choose from.
    - Rule destinations: Select destinations and destination groups to be notified when the event's condition occurs. Here, you can select one or more existing destinations or also define a new destination (see Figure 7-47).

Figure 7-47 Select destination

- Rule snooze: Defines whether the system repeatedly alerts the defined destination until the event is cleared. If so, a snooze time must be selected. Check **Use snooze timer** and enter a snooze time in minutes.
- Rule escalation: Allows the system to send alerts by other rules if the event is not cleared within a certain time. If so, an escalation time and rule must be specified:
  i. Check **Use escalation rule**.
  ii. Click **Escalation Rule**.
  iii. Enter an escalation time in minutes.
  iv. Click **Create Escalation Rule**.

12. On the summary window shown in Figure 7-48, review the information that you entered. Go back if you need to make changes. If everything is correct, click **Create**.

Figure 7-48 Rule Create window

Setting up notifications and rules using the XCLI

You use the same process to set up the XIV Storage System for notification using the XCLI as you used with the XIV Storage Management GUI. The three-step process includes all the required configurations to allow the XIV Storage System to provide notification of events:

- Gateway
- Destination
- Rules

The gateway definition is used for SMTP and SMS messages. There are various commands that are used to create and manage the gateways for the XIV Storage System. Example 7-25 shows an SMTP gateway being defined. The gateway is named *test* and the messages from the XIV Storage System are addressed to `xiv@us.ibm.com`.

When added, the existing gateways are listed for confirmation. In addition to gateway address and sender address, the port and reply-to address can also be specified. There are various other commands that are available for managing a gateway.

Example 7-25 The smtpgw_define command

```
>> smtpgw_define smtpgw=test address=test.ibm.com from_address=xiv@us.ibm.com
Command executed successfully.

>> smtpgw_list
Name                Address         Priority
ITSO Mail Gateway   us.ibm.com      1
test                test.ibm.com    2
```

The SMS gateway is defined in a similar method. The difference is that the fields can use tokens to create variable text instead of static text. When specifying the address to send the SMS message, tokens can be used instead of hardcoded values. In addition, the message body also uses a token to have the error message sent instead of a hardcoded text.

Gateways: An SMTP gateway must be defined before you can define an SMS gateway because SMS messages are sent from the XIV Storage System in the form of an email.

Example 7-26 provides an example of defining an SMS gateway. The following tokens are available to be used for the SMS gateway definition:

- {*areacode*}: This escape sequence is replaced by the destination's mobile or cellular phone number area code.
- {*number*}: This escape sequence is replaced by the destination's cellular local number.
- {*message*}: This escape sequence is replaced by the text to be shown to the user.
- \{, \}, \\: These symbols are replaced by the {, } (for \{, \}) or \ (for \\).

Example 7-26 The smsgw_define command

```
>> smsgw_define smsgw=test email_address={areacode}{number}@smstest.ibm.com
subject_line="XIV System Event Notification" email_body={message}
Command executed successfully.
>> smsgw_list
Name  Email Address                        SMTP Gateways
test  {areacode}{number}@smstest.ibm.com  all
```

When the gateways are defined, the destination settings can be defined. There are three types of destinations:

- SMTP or email
- SMS
- SNMP

Example 7-27 provides an example of creating a destination for all three types of notifications. For the email notification, the destination receives a test message every Monday at 12:00. Each destination can be set to receive notifications on multiple days of the week at multiple times.

Example 7-27 Destination definitions

```
>> dest_define dest=emailtest type=EMAIL email_address=test@ibm.com smtpgws=ALL
heartbeat_test_hour=12:00 heartbeat_test_days=Mon
Command executed successfully.

>> dest_define dest=smstest type=SMS area_code=555 number=5555555 smsgws=ALL
Command executed successfully.

>> dest_define dest=snmptest type=SNMP snmp_manager=9.9.9.9
Command executed successfully.

>> dest_list
Name           Type   Email Address  Area Code Phone Number SNMP Manager            User
ITSO_Catcher   SNMP                                         itsocatcher.us.ibm.com
smstest        SMS                     555      5555555
snmptest       SNMP                                         9.9.9.9
emailtest      EMAIL  test@ibm.com
```

Finally, the rules can be set for which messages can be sent. Example 7-28 provides two examples of setting up rules. The first rule is for SNMP and email messages, and all messages, even informational messages, are sent to the processing servers. The second example creates a rule for SMS messages. Only critical messages are sent to the SMS server, and they are sent every 15 minutes until the error condition is cleared.

Example 7-28 Rule definitions

```
>> rule_create rule=emailtest min_severity=informational dests=emailtest,snmptest
Command executed successfully.

>>rule_create rule=smstest min_severity=critical dests=smstest snooze_time=15
Command executed successfully.

>> rule_list
Name         Minimum Severity   Event Codes   Except Codes   Destinations         Active   Escalation Only
ITSO_Major   Major              all                          ITSO_Catcher         yes      no
emailtest    Informational      all                          emailtest,snmptest   yes      no
smstest      Critical           all                          smstest              yes      no
```

Example 7-29 shows how to delete rules, destinations, and gateways. It is not possible to delete a destination if a rule is using that destination. And, it is not possible to delete a gateway if a destination is pointing to that gateway.

Example 7-29 Deletion of notification setup

```
>> rule_delete -y rule=smstest
```

```
Command executed successfully.
>> dest_delete -y dest=smstest
Command executed successfully.
>> smsgw_delete -y smsgw=test
Command executed successfully.
```

7.4 Mobile push notification

The XIV Mobile Notification Service allows client to receive real-time alerts about major or critical events from the IBM XIV Service Center anytime, anywhere. This enables them to react rapidly to potential storage issues.The mobile device notifications feature was introduced with XIV Storage System Software v11.5.

The prerequisite is to have Call Home enabled on the XIV system (see 7.5.1, “Call Home feature” on page 372). The IBM XIV Mobile Dashboard application must be to be installed on a mobile device. The Mobile Dashboard application can be downloaded from the following mobile device applications store:

- Android device store:

  https://play.google.com/store/apps/details?id=com.ibm.xiv.mobile.android&hl=en
- iOS device store

  http://itunes.apple.com/us/app/ibm-xiv-mobile-dashboard/id465595012?mt=8

The subscription to the push service is on a user-system basis. On an iOS platform, after installing and signing on to the mobile application for the first time, you must confirm to accept push notifications. On an Android platform, users are automatically enabled for notifications.

A new Notifications setting in the mobile user controls panel lets you specify whether or not to receive notifications from the XIV system; See Figure 7-49.

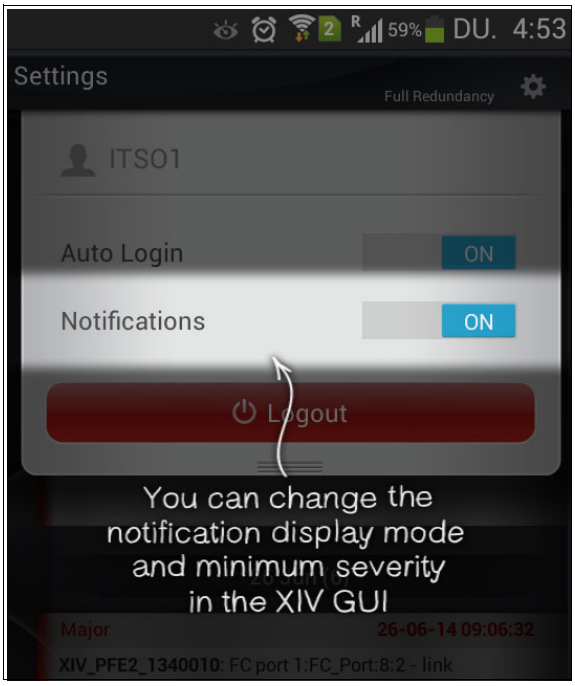


Figure 7-49 Controlling notification on mobile device

The push notification service is based on the XIV Service Center. When new events are received from XIV system and match with certain event rule setup in Service Center, the relevant notification is pushed via Notification Service Provider (Apple or Google) to the registered mobile client as shown in Figure 7-50.

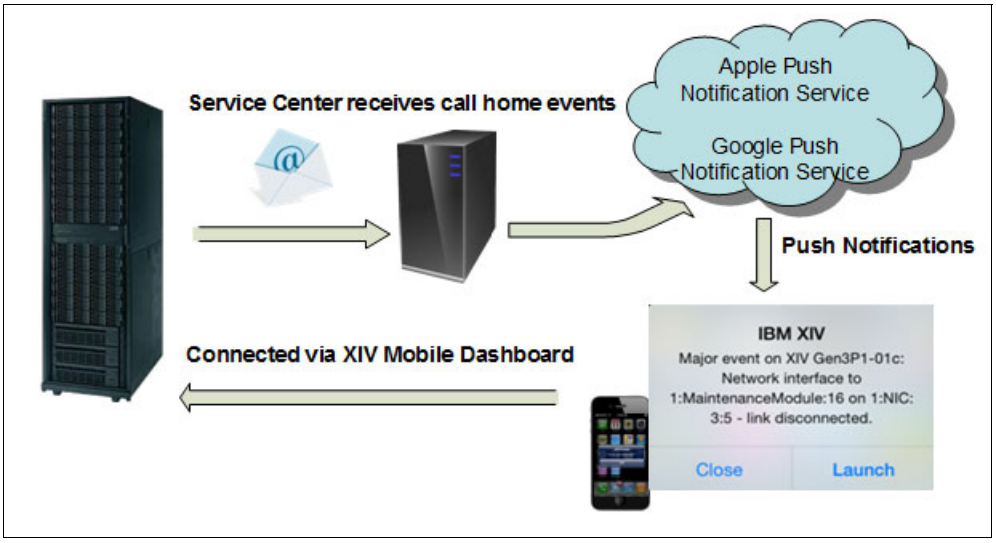

Service Center receives call home events
Apple Push Notification Service
Google Push Notification Service
Push Notifications
IBM XIV
Major event on XIV Gen3P1-01c: Network interface to 1:MaintenanceModule:16 on 1:NIC: 3:5 - link disconnected.
Close
Launch
Connected via XIV Mobile Dashboard

Figure 7-50 Push Notification process flow

An auto-login feature allows users to tap on the notification event and automatically log in to the system. These system settings can be edited at any time.

After you are logged in to the system, the Events view shows all relevant notifications as set by the demand minimum severity filter. Note that this events view is a new element (the Events icon is at the bottom) in Version 1.4 of the IBM XIV mobile dashboard, as illustrated in Figure 7-51. (This screen capture is based on an early version of the code, so the rendering might be somewhat different in the final version.)

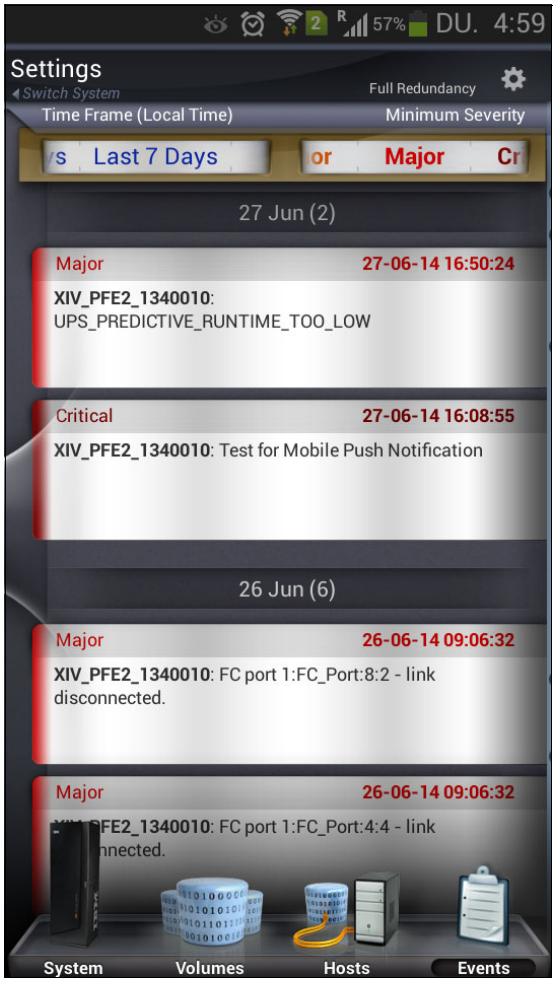


Figure 7-51 Displaying the Events view per notifications

A storage administrator can control how much information and configure which notifications are sent to a mobile device. Navigate to **Access** → **Mobile Notifications** from the XIV Storage Management GUI as shown in Figure 7-52.

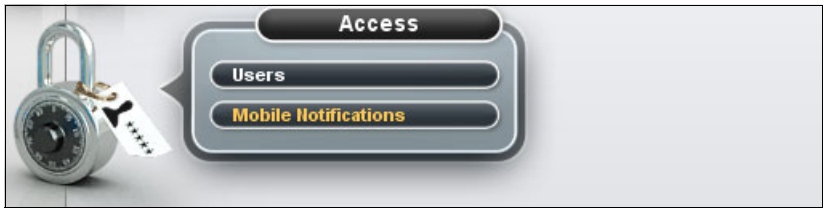


Figure 7-52 Select Mobile Notifications

Notifications can entail informative event descriptions and are configurable (see Figure 7-53). By default, major and critical issues are sent in preview-only mode.

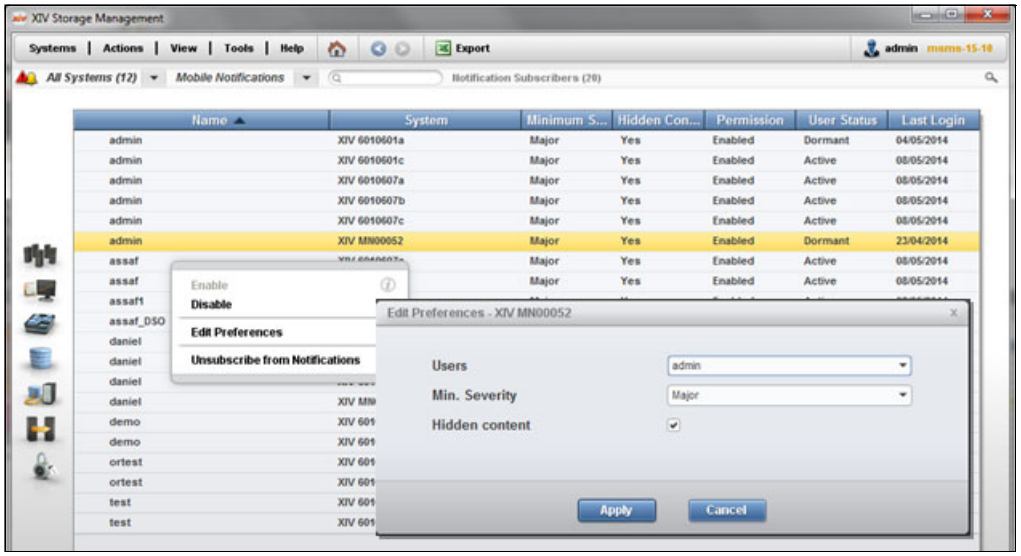


Figure 7-53 Notification control

User permissions and status are managed from the same panel in the IBM XIV Storage Management GUI:

- Active user

  An active user receives notifications based on the defined preferences.

- Dormant user

  If a user has not used the service for over a week, the service is considered to be dormant. In this case, the user receives push notifications, but with no content, regardless of the defined preferences.

- Non-active user

  If a user has not logged in for more than 30 days, the user subscription is considered inactive. Logging out from the XIV Mobile application disables push notifications. Notifications can also be disabled from the settings window.

If a mobile user unsubscribes from receiving push notifications, the user registration remains on the XIV system. To completely remove the user privileges, the user must be removed from the system via the GUI.

7.5 Call Home and remote support

The Call Home function allows the XIV Storage System to send event notifications to the XIV Remote Support Center (XRSC). This function enables both proactive and failure notifications to be sent directly to IBM for analysis.

The XRSC takes appropriate action, up to dispatching an IBM Service Support Representative (SSR) with a replacement part or engaging Level 2 or higher-level support to ensure complete problem determination and a solution.

Important: The configuration of Call Home and the configuration of remote support facilities are recommended to assist with failure detection, diagnosis, and resolution.

7.5.1 Call Home feature

Call Home is always configured to use SMTP, and is only configured by qualified IBM SSRs, typically when the XIV Storage System is first installed.

If an event is received by the XRSC that requires service or investigation, the event typically triggers a new IBM problem management record (PMR).

Because Call Home uses the client's network and SMTP service, IBM cannot guarantee the delivery of events. Therefore, events must be monitored by the client as described in 7.3, "XIV Storage System event notification" on page 359.

When the Call Home feature is configured and events are received by the XRSC, periodic heartbeat events are also received. The heartbeats are monitored by IBM and the client is notified if the heartbeats are not received or are no longer received.

There is an IBM service, called Electronic Service Call (ESC+), where you can monitor service calls that are specific to a customer account ID. The ESC+ gives you the ability to verify that a service ticket has been raised following an appropriate event, and to open a new service call. The ESC+ website is shown:

`http://www.ibm.com/support/esc`

Contact your local IBM SSR to have your user ID associated with a specific customer ID in ESC+.

The SMTP address for Call Home is configured separately from the general XIV Storage System SMTP setting. If the client's mail server gateway changes, a service call must be logged to have the internal Call Home SMTP setting changed.

Tip: Email relaying on the SMTP gateway server might need to be enabled to allow Call Home events to be sent to IBM.

7.5.2 Remote support

The XIV Storage System is repaired by trained IBM service personnel, either remotely with the help of the XRSC, or on site by an IBM SSR. When problems arise, a remote support specialist can connect to the system to analyze the problem, repair it remotely if possible, or assist the IBM SSR who is onsite.

Remote support has three ways to connect the system. Depending on the client's choice, the support specialist can connect by one of the following methods:

- Using a modem dial-up connection through an analog phone line provided by the client
- Using a secure, high-speed direct connection through the Internet to the XIV Storage System
- Using the XRSC, which allows the client to initiate a secure connection from the XIV Storage System to IBM. Using XRSC, the XIV Storage System makes a connection to an external XRSC server. Using an internal XRSC server, the XRSC can connect to the XIV Storage System through the connection made to the external server. For more details, see Figure 7-54.

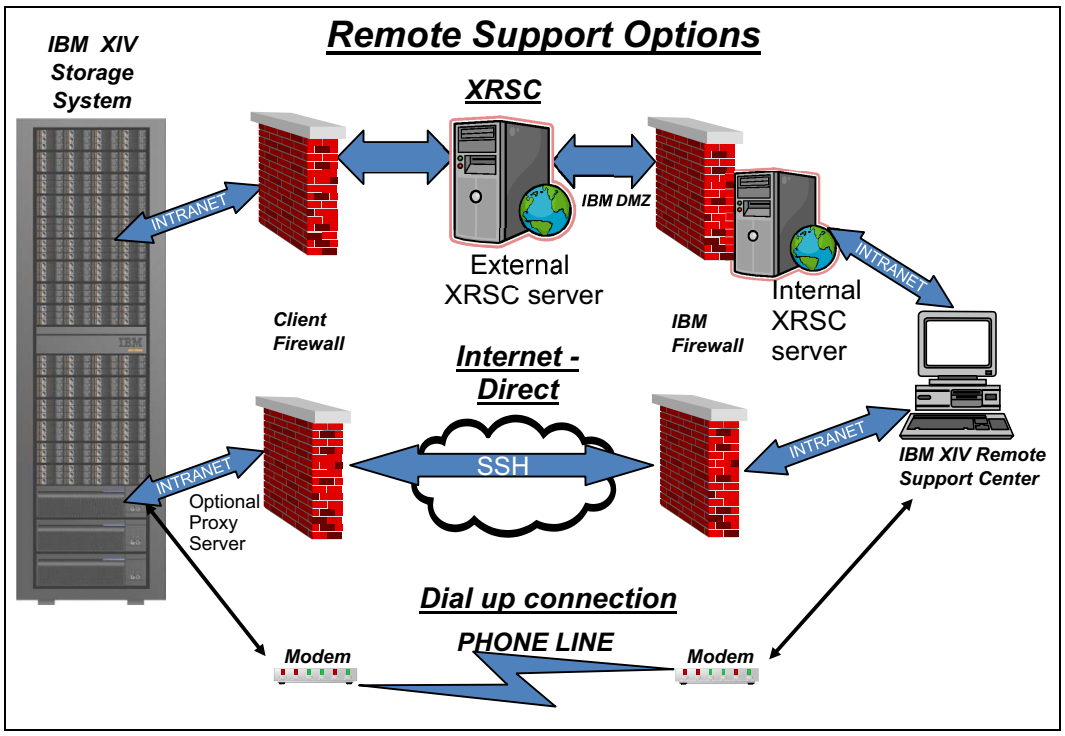


Figure 7-54 Remote support connections

XRSC: We encourage all clients to use the secure, high-speed remote support solution enabled by the XRSC.

These possibilities are shown in Figure 7-54. If there are problems, the remote specialist is able to analyze problems and also assist an IBM SSR dispatched onsite in repairing the system or in replacing field-replaceable units (FRUs).

To enable remote support, you must allow an external connection, such as either one of the following methods:

- A telephone line
- An Internet connection through your firewall that allows IBM to use a Secure Shell (SSH) connection to your XIV Storage System

XRSC connection

The XRSC uses a high-speed Internet connection, but it gives the client the ability to initiate an outbound SSH call to a secure IBM server.

Firewall rules might need to be configured at the client firewall to allow the XIV Storage System VPN/Management ports to connect to the XRSC.

Tip: The type of access required for a remote support connection is “outbound port 22/ssh” from the XIV Storage System network ports.

The XRSC consists of the XIV Storage System’s internal functions with a set of globally deployed supporting servers. Together, they provide secure IBM support access to the XIV Storage System when necessary and when authorized by the client’s personnel.

Objectives and components

The XRSC meets three major objectives:

- Security from malicious attack:

  The XRSC was designed with security as a major concern, while keeping the system architecture simple and easy to deploy. It relies on standard, proven technologies and minimizes the logic (code) that must be located either on the external XRSC server or on client machines. The architecture of the XRSC anticipates and handles various potential attack vectors, including encryption protocol and authentication logic. The design of the XRSC identifies, analyzes, and defends against these and other possible attack routes.

- Simplicity of implementation and design:

  Simplicity enhances security, facilitates implementation at the client site, decreases staff training costs, and ultimately lowers total cost of ownership (TCO) by allowing more issues to be resolved quickly and with fewer unnecessary visits by support personnel to client premises.

- Accessibility by support personnel:

  The auditing features of the XRSC both enhance security and enable easy reassignment of a support case among support personnel. The XRSC can expedite escalation of support cases to higher-level support and allows IBM SSRs to pool knowledge to resolve complex client issues.

Underlying architecture

The XIV remote support mechanism has four major components:

- Remote Support Client (machine internal)

  The Remote Support Client is a software component inside the XIV Storage System that handles remote support connectivity. It relies only on a single outgoing Transmission Control Protocol (TCP) connection, and has no capability to receive inbound connections of any kind. The Client is controlled by using XCLI and is charged with starting a connection, terminating a connection (because of timeout or client request), and trying the connection again in case it terminates unexpectedly.

- Optional Remote Support Proxy

  The Remote Support Client can access the Remote Support Center Front Server directly, or through an optional proxy server. The optional Remote Support Proxy can be used when one or more IBM XIV systems do not have direct access to the Internet (for example, because of firewall restrictions). You can use the Remote Support Proxy to facilitate the connection to the XRSC. More information about the Remote Support Proxy can be found in the *IBM XIV Remote Support Proxy Installation and User's Guide*, GA32-0795.

- Remote Support Center Front Server (Internet)

  Front Servers are on an IBM DMZ of the Internet and receive connections from the Remote Support Client and the IBM XIV Remote Support Back Server. Front Servers are security-hardened machines that provide a minimal set of services, namely, maintaining connectivity to connected Clients and to the Back Server. They are strictly inbound, and never initiate anything on their own accord. No sensitive information is ever stored on the Front Server, and all data passing through the Front Server from the Client to the Back Server is encrypted so that the Front Server cannot access this data.

- Remote Support Center Back Server (IBM intranet)

  The Back Server manages most of the logic of the system. It is located within the IBM intranet. The Back Server is access controlled. Only IBM employees authorized to perform remote support of the XIV Storage System are allowed to use it, and only through specific

support interfaces, not with a CLI or a GUI shell. The Back Server is in charge of authenticating a support person. It provides the support person with a user interface (UI) through which to choose a system to support based on the support person's permissions. It also provides the list of systems currently connected to the Front Servers and manages the remote support session as it progresses (logging it, allowing additional support persons to join the session, and so on). The Back Server maintains connection to all Front Servers. Support people connect to the Back Server using any SSH client or an HTTPS connection with any browser.

Figure 7-55 provides a representation of the data flow of the XIV Storage System to IBM Support.

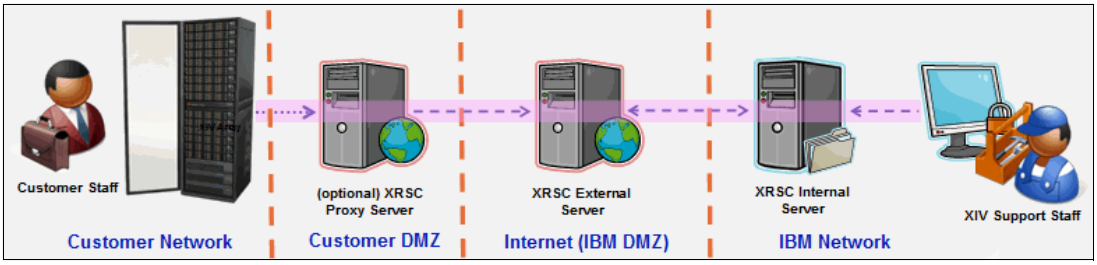


Figure 7-55 XRSC

To initiate the remote connection process, the following steps are performed:

1. The client initiates an Internet-based SSH connection to XRSC either through the XIV Storage Management GUI or XCLI.
2. XRSC identifies the XIV Storage System and marks it as “connected”.
3. Support personnel connect to XRSC using SSH over the IBM intranet.
4. XRSC authenticates the support person against the IBM intranet.
5. XRSC then shows the connected client system available to the support personnel.
6. The IBM Support person then chooses which system to support and to which system to connect:
   - Only permitted IBM XIV systems are shown.
   - IBM Support personnel log their intended activity.
7. A fully recorded support session commences.
8. When complete, the support person terminates the session and the XRSC disconnects the XIV Storage System from the remote support system.

Client use scenario

After a client has decided that remote support is needed, possibly because of a telephone request from a remote support professional, communication between the XRSC and the client site must be established. This process is always initiated by the client and always remains under the client’s control. It proceeds as follows:

- The client tells the XIV Storage System to connect to the XRSC using a specific XCLI command. Until this command is issued, an XIV Storage System will not connect to the XRSC for any reason.

- The client can set a regular timeout and an “idle timeout” for the connection, using an optional parameter in the connection command. A *timeout* causes the connection to be dropped after a specified length of time whether a support session is in progress or not, whereas an *idle timeout* causes the connection to be dropped after a certain period of inactivity. Specifying the keyword “never” for these two parameters allows the connection to remain established indefinitely.
- The client can also choose to set a unique password for a session when it is initiated. This password must be provided to the IBM Support representative for them to complete the authentication process and access the XIV Storage System.
- In addition, the client can see the XRSC connectivity status at any time (disconnected, idle, or in-session) and can choose to forcibly disconnect the connection regardless of whether a support session is in progress or not.

Although forcibly terminating a support session using the timeout or force-disconnect mechanisms is not desirable, these mechanisms do exist and the client has full control over whether to proceed with a support session.

The initiation or loss of connection to the XRSC causes the system to generate events that can be seen on the machine's event log. These events can be forwarded to any destination of the client’s choice (like any other event using the XIV Storage System event-rules mechanism). Connection loss events, whether because of a transport error, timeout, or due to specific client action, specify whether the connection was idle or in-use at the time of disconnection. A warning event is issued 15 minutes before the timeout parameter disconnects a busy support session. If the connection is lost because of a network error before the timeout for the session has expired, the system automatically tries to reconnect to any of the configured XRSC servers.

While a support session is in progress, the XIV Storage System generates events and shows the machine’s status on the GUI window as usual. Therefore, for example, the client can see the process of phasing out a module or the restart of client-visible machine services as they happen.

Starting an XRSC connection using the GUI

The XRSC definitions are configured at the time the system is installed. To start the connection from the GUI, click **Tools** → **Support Center**, as shown in Figure 7-56.

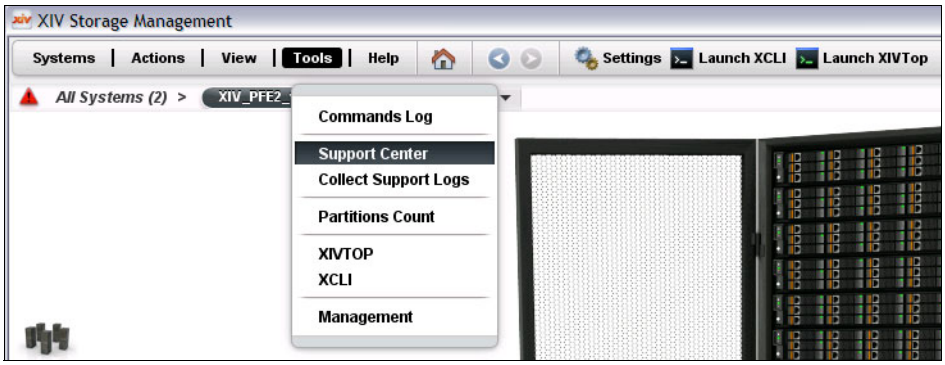


Figure 7-56 Starting an XRSC Remote Support Center connection

Select the wanted support center from the list and click the **Connect Support Center** icon, as shown in Figure 7-57.

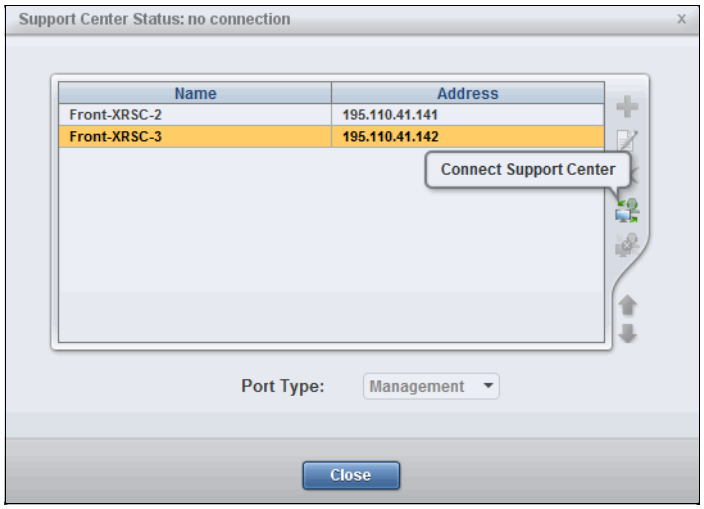


Figure 7-57 Select the wanted support center and click the Connect Support Center icon

The window shown in Figure 7-58 prompts you for the Session Timeout, Idle Timeout, and Connection Password values. The timeout values are specified in minutes and disconnect the session when they expire. The password, if set, must be given to the IBM Support representative for them to establish a connection. Use of a password is not required and can be left blank.

Using **Never** for the timeout values results in the connection remaining open until explicitly disconnected.

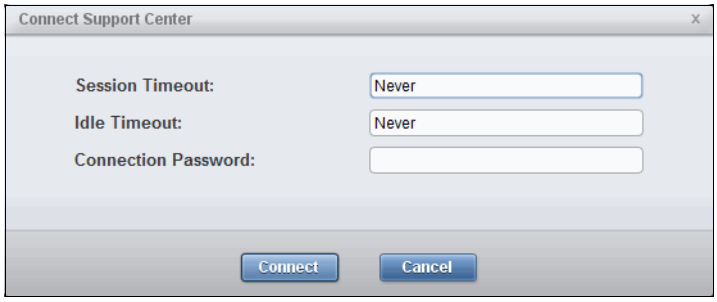


Figure 7-58 Starting an XRSC connection

After the connection to the XRSC has been established, you can disconnect the session by clicking the **Disconnect Support Center** icon, as shown in Figure 7-59 on page 378.

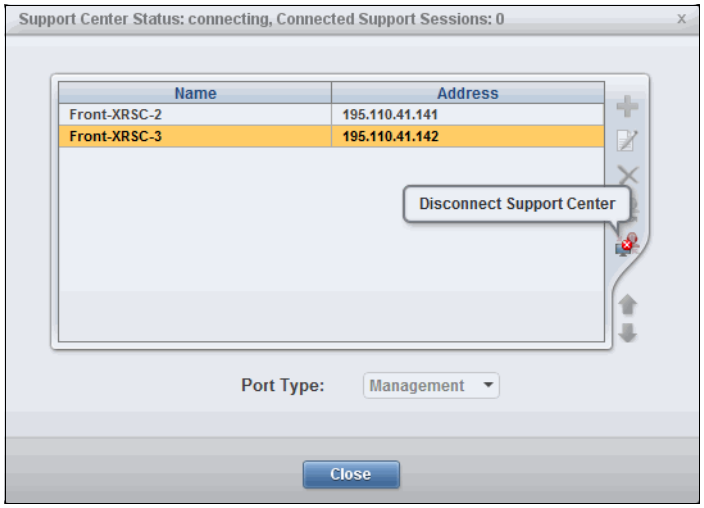


Figure 7-59 Disconnecting a Remote Support Center session

Starting an XRSC connection using XCLI

The following XCLI commands are used to manage Remote Support Center connections:

- **support_center_list**
- **support_center_status**
- **support_center_connect**
- **support_center_disconnect**

To start an XRSC connection using GUI, open a list of the available support centers that were configured during system installation by issuing the support_center_list XCLI command, as shown in Example 7-30.

Example 7-30 List available support centers

```
>>support_center_list
Name           Address          Port   Priority
Front-XRSC-2   195.110.41.141   22     0
Front-XRSC-3   195.110.41.142   22     0
```

Contact your support representative if the **support_center_list** commands do not return a value and you want to use this feature.

Next, you can see what the status of the connection is by running the **support_center_status** command (Example 7-31).

Example 7-31 Check support center connection status

```
>>support_center_status
State          Connected sessions  Timeout (min)  Module  Connected since
no connection
```

In this example, no connections are currently established. To start a session, run **support_center_connect** followed by **support_center_status**, as shown in Example 7-32 on page 379.

Example 7-32 Connect to support center

```
>> support_center_connect
command 0:
administrator:
    command:
        code = "SUCCESS"
        status = "0"
        status_str = "Command completed successfully"
aserver = "DELIVERY_SUCCESSFUL"

>> support_center_status
State   Connected sessions   Timeout (min)   Module       Connected since
------- -------------------- --------------- ------------ --------------------
idle    0                    no timeout      1:Module:4   2010-10-08 10:45:35
```

If no parameters are specified with the **support_center_connect** command, the timeout values default to never and no password is used. You can specify the parameters for **timeout**, **idle_timeout**, and **password**, as shown in Example 7-33. The timeout values are in the format *hh*:*mm* when using the XCLI commands.

Example 7-33 Specify connection parameters

```
>> support_center_connect timeout=01:00 idle_timeout=00:30 password=4support
command 0:
administrator:
    command:
        code = "SUCCESS"
        status = "0"
        status_str = "Command completed successfully"
aserver = "DELIVERY_SUCCESSFUL"

>> support_center_status
State   Connected sessions   Timeout (min)   Module       Connected since
------- -------------------- --------------- ------------ --------------------
idle    0                    27.6            1:Module:4   2010-10-08 10:49:40
```

The status shows an idle state until an XRSC representative establishes a connection, at which time it shows a state of busy, as shown in Example 7-34.

Example 7-34 Connection busy

```
>> support_center_status
State   Connected sessions   Timeout (min)   Module       Connected since
------- -------------------- --------------- ------------ --------------------
busy    1                    40.2            1:Module:4   2010-10-08 10:49:40
```

To terminate the Remote Support Center connection, run the **support_center_disconnect** command, as shown in Example 7-35.

Example 7-35 Disconnect from the support center

```
>> support_center_disconnect
command 0:
administrator:
    command:
        code = "SUCCESS"
        status = "0"
        status_str = "Command completed successfully"
aserver = "DELIVERY_SUCCESSFUL"

>> support_center_status
State           Connected sessions   Timeout (min)   Module   Connected since
--------------- -------------------- --------------- -------- -----------------
no connection   ?                    ?               ?        ?
```

7.5.3 XIV Remote Support Proxy

The optional XIV Remote Support Proxy agent can be used when one or more XIV systems do not have direct access to the Internet (for example, due to firewall restrictions). You can use the XIV Remote Support Proxy agent to facilitate the connection to the XRSC.

After the XIV Remote Support Proxy agent is configured, the connection to the XRSC is performed normally from the XIV Storage System, as described in “XRSC connection” on page 373.

The agent is a small program that runs on the following Linux versions:

- Red Hat Enterprise Linux, Version 6.0 or later, for x86 and x86-64 systems
- Red Hat Enterprise Linux, Version 5.1 or later, for x86 and x86-64 systems
- Red Hat Enterprise Linux, Version 4.6 or later, for x86 and x86-64 systems
- SUSE Linux Enterprise Server 11 or later, for x86 and x86-64 systems

Important: Review the latest release notes for updates.

The host running the agent must have TCP/443 outbound access to XRSC addresses (information supplied by IBM Support) and listens for inbound connections from the IBM XIV systems.

Figure 7-60 shows the necessary ports.

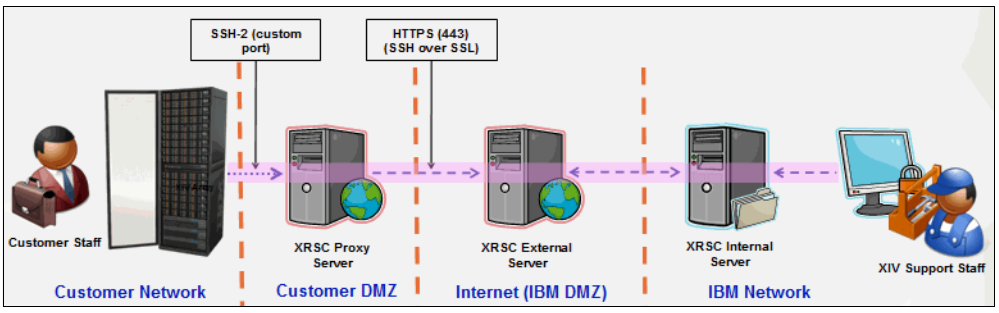


Figure 7-60 XIV Remote Support Proxy agent ports

For more information about the XIV Remote Support Proxy, see the IBM XIV Storage System documentation in the IBM Knowledge Center:

http://ibm.co/1Os4BzX

7.5.4 Installation

The installation files and documentation are at the storage portal website for XIV:

http://ibm.co/1xltw1V

On this website, expand the list by clicking the **More Results** link.

After you download the correct package to the Linux host, you can run the file as root and it starts the installation wizard.

Example 7-36 shows the installation of the proxy agent.

Example 7-36 Remote Support Proxy agent installation

```
# ./xivproxy-installer-rpm-1.3.1-b612.sles11.x86_64.bin
Starting installer, please wait...
International License Agreement for Non-Warranted Programs

Part 1 - General Terms

BY DOWNLOADING, INSTALLING, COPYING, ACCESSING, CLICKING ON AN "ACCEPT" BUTTON, OR
OTHERWISE USING THE PROGRAM, LICENSEE AGREES TO THE TERMS OF THIS AGREEMENT. IF
YOU ARE ACCEPTING THESE TERMS ON BEHALF OF LICENSEE, YOU REPRESENT AND WARRANT
THAT YOU HAVE FULL AUTHORITY TO BIND LICENSEE TO THESE TERMS. IF YOU DO NOT AGREE
TO THESE TERMS,

Press Enter to continue viewing the license agreement, or, Enter "1" to accept the
agreement, "2" to decline it or "99" to go back to the previous screen, "3" Print.
1
xivproxy                  0:off  1:off  2:off  3:on   4:off  5:on   6:off
Installation completed successfully.
```

You can edit the `proxy.conf` file, which is in the `/etc/xiv/` folder, to add the relevant connectivity settings, as shown in Example 7-37.

Example 7-37 Extract of the proxy.conf file

```
# Configuration file for XIV remote support proxy
...
# Network interface and port that the XIV system will connect to
ListenInterface eth0
ListenPort 8988

# Remote support server address as supplied by IBM
TargetAddress 195.110.41.141
TargetPort 443

# Network interface (Bo for local) for status queries
StatusInterface eth0
StatusPort 8989
```

```
# HTTP proxy for connecting to the Internet
# HTTPProxyHost
# HTTPProxyPort
...
```

The parameters referenced in the `proxy.conf` file include the following parameters:

`ListenInterface` The network interface name on the host running the Remote Proxy agent that the XIV Storage System contacts when initiating a Remote support connection (for example, eth0).

`ListenPort` The TCP port on which `ListenIinterface` listens (for example, 8988).

Important: The XIV Storage System must be able to contact the host running the XIV Remote Support Proxy agent on the specified `ListenInterface` and `ListenPort` to run a remote support connection.

`TargetAddress` The network address of the XRSC server to which the XIV Remote Support Proxy agent initiates a connection. IBM supplies the address.

`TargetPort` The TCP port of the XRSC server. This port is normally set to 443, but confirm this port when you get the `TargetAddress` information from IBM Support.

`StatusInterface` The network interface name where you can query the status of the XIV Remote Support Proxy agent and show any active remote support connections.

`StatusPort` The TCP port on which `StatusIinterface` listens (for example, 8989).

`HTTPProxyHost` The network address of an external web proxy. This parameter is an optional parameter if there is an existing web proxy service and it is not possible to open the firewall for the host running the proxy.

`HTTPProxyPort` The TCP port on which `HTTPProxyHost` listens.

After you configure the `proxy.conf` file, you can start the proxy service by running **`service xivproxy start`** and then run **`service xivproxy status`** to confirm that the XIV Remote Support Proxy agent is running, as shown in Example 7-38.

Example 7-38 Starting the Remote Proxy and check its status

```
bc-h-15-b6:/ # service xivproxy start
Starting IBM XIV remote support proxy:
done
bc-h-15-b6:/ # service xivproxy status
IBM XIV remote support proxy running
Listen address    : 9.155.113.137:8988
Target address    : 195.110.41.141:443
Running since     : Sep-27 15:40:56
Open connections  : 0
Failed connections: 0
Total connections : 0
```

Example 7-39 on page 383 shows a query of the XIV Remote Support Proxy agent to check its status from a remote location by using Telnet. In this example, there are two active remote support connections, from separate IBM XIV systems, using the proxy.

Example 7-39 Querying the status of the XIV Remote Support Proxy agent using Telnet

```
# telnet 9.155.113.137 8989
Trying 9.155.113.137...
Connected to 9.155.113.137.
Escape character is '^]'.
Listen address    : 9.155.113.137:8988
Target address    : 195.110.41.141:443
Running since     : Sep-27 15:40:56
Open connections  : 2
Failed connections: 0
Total connections : 2
Connection to 9.155.113.137 closed by foreign host.
```

After the XIV Remote Support Proxy agent has been configured and started successfully, new **support_center** definitions must be entered on the XIV Storage System. These definitions are done by your XIV Storage System support personnel. An example of a configuration is shown in Example 7-40.

Example 7-40 support_center_list of entry using XIV Remote Proxy

```
XIV PFE-GEN3-1310133>>support_center_list
Name       Address         Port   Priority
xivproxy   9.155.113.137   8988   5
```

Now, a remote support connection can be made, as shown in Example 7-41.

Example 7-41 Establishing a remote support connection

```
XIV PFE-GEN3-1310133>>support_center_connect
Command executed successfully.
command_line=/xiv/python/nextra/remote_support_client/remote_support_client connect -I 2810-114-1310133 -m
1:Module:2 -d xivproxy:9.155.113.137:8988:5

XIV PFE-GEN3-1310133>>support_center_status
State  Connected sessions  Timeout (min)  Module      Connected since      Destination
idle   0                   no timeout     1:Module:2  2011-09-27 15:55:48  xivproxy(9.155.113.137:8988)
```

More information about the Remote Support Proxy is in the *IBM XIV Remote Support Proxy Installation and User's Guide*, GA32-0795.

7.6 Simple Network Management Protocol-based monitoring

The XIV Storage System supports Simple Network Management Protocol (SNMP) for monitoring.

7.6.1 Monitoring tools

SNMP-based monitoring tools, such as IBM Tivoli NetView® or IBM Systems Director, can be used to monitor the XIV Storage System.

Simple Network Management Protocol

SNMP is an industry-standard set of functions for monitoring and managing TCP/IP-based networks and systems. SNMP includes a protocol, a database specification, and a set of data objects. A set of data objects forms a Management Information Base (MIB).

The SNMP protocol defines two terms, *agent* and *manager*, instead of the client and server terms that are used in many other TCP/IP protocols. An SNMP agent is implemented in the XIV Storage System, which sends SNMP traps to an SNMP manager (such as IBM Systems Director) to indicate that an event has occurred. By default, the trap is sent to UDP port 162. The SNMP manager can also request certain information from the XIV Storage System using SNMP **get** or **walk** commands. These commands are sent to the XIV Storage System on UDP port 161.

Management Information Base

The information that you can receive from the XIV Storage System by sending an SNMP **get** request, or the contents of an SNMP trap sent from the XIV Storage System, are defined in an MIB. The structure of a MIB is defined as an Internet standard in RFC 1155. The MIB forms a tree structure that consists of object identifiers (OIDs).

Most hardware and software vendors provide you with extended MIB objects to support their own requirements. The SNMP standards allow this extension by using the private subtree, which is called an *enterprise-specific MIB*. Because each vendor has a unique MIB subtree under the private subtree, there is no conflict among vendors' original MIB extensions.

You can use the **mib_get** command from the XCLI.

For an SNMP trap sent by an XIV Storage System, the MIB defines the following object IDs:

```
1.3.6.1.4.1.2021.77.1.3.1.1.1 xivEventIndex A unique value for each event
1.3.6.1.4.1.2021.77.1.3.1.1.2 xivEventCode The code of the event
1.3.6.1.4.1.2021.77.1.3.1.1.3 xivEventTime The time of the event
1.3.6.1.4.1.2021.77.1.3.1.1.4 xivEventDescription A description of the event
1.3.6.1.4.1.2021.77.1.3.1.1.5 xivEventSeverity The severity of the event
1.3.6.1.4.1.2021.77.1.3.1.1.6 xivEventTroubleshooting Troubleshooting information
```

XIV Storage System SNMP setup: Configuring a new trap destination

To use SNMP monitoring with the XIV Storage System, you must first configure it to send SNMP traps to an SNMP manager (such as the IBM Systems Director server), which is defined in your environment. Figure 7-61 on page 385 shows where to start to set up the SNMP destination. For more information, see "Setup notification and rules with the GUI" on page 359.

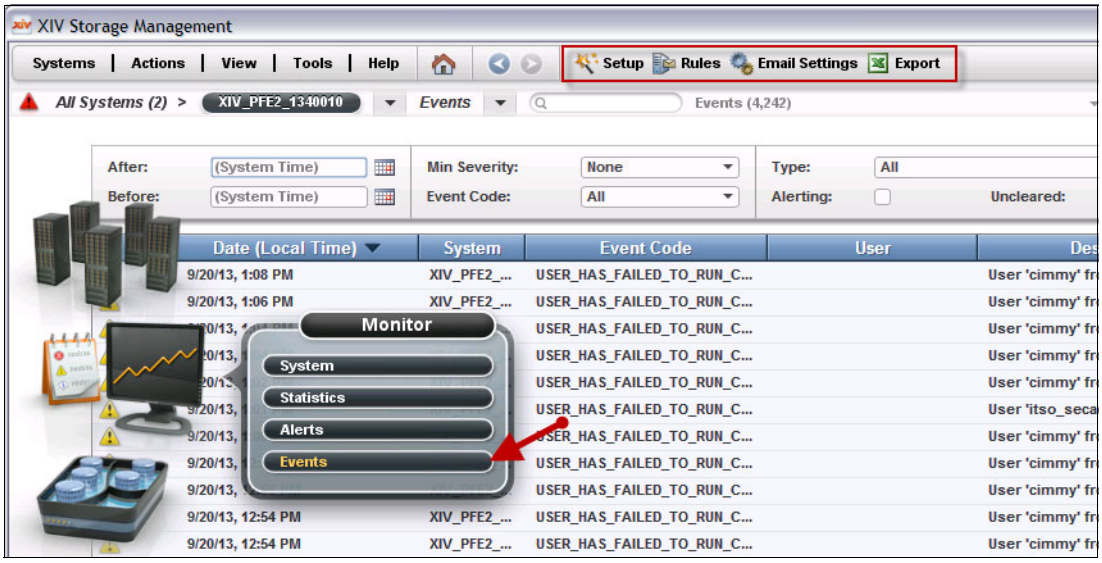


Figure 7-61 Configure destination

To configure a new SNMP destination, complete the following steps:

1. From the XIV Storage Management GUI main window, select the **Monitor** icon.
2. From the Monitor menu, select **Events** to open the Events window, as shown in Figure 7-61.

   From the toolbar, select **Rules** and perform the following steps:

   a. Click **Destinations**. The Destinations window opens.

   b. Select **SNMP** from the Destinations drop-down list.

   c. Click the green plus sign (**+**) and click **Destination** to add a destination, as shown in Figure 7-62.

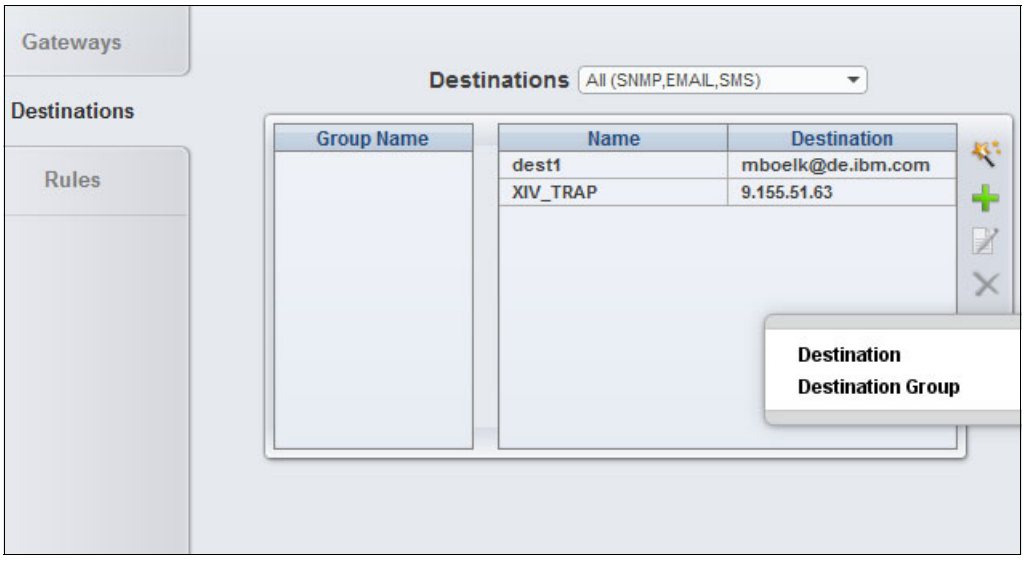


Figure 7-62 Add SNMP destination

3. From the Define Destination window, which is now open, enter a Domain name if any specified (default is no domain), a Destination Name (a unique name of your choice) and

the IP or Domain Name System (DNS) of the server where the SNMP Management software is installed. See Figure 7-63.

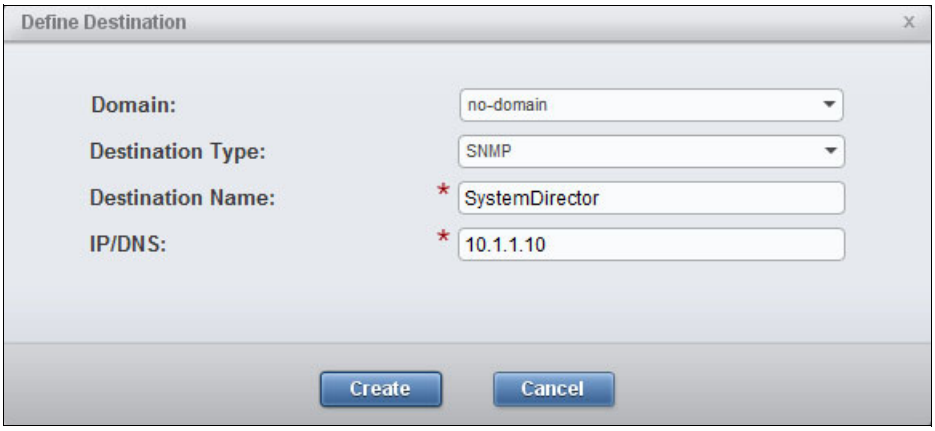


Figure 7-63 Define SNMP destination

4. Click **Define** to effectively add the SNMP Manager as a destination for SNMP traps.

In addition, set up the rules for the defined SNMP destination, as described in "Setup notification and rules with the GUI" on page 359. Afterward, the XIV Storage System is set up to send SNMP traps to the defined SNMP manager. The SNMP Manager software processes the received information (SNMP traps) according to the MIB file.

7.6.2 Using SNMP commands to confirm the XIV Storage System status

Although SNMP traps can be received from the XIV Storage System, you can also send SNMP **`get`** or **`walk`** commands to collect status information from the XIV Storage System. To accomplish this task, you must use an SNMP manager that supports this task and you need to import the XIV Storage System MIB into that manager.

To send SNMP **`get`** commands, you must know the SNMP community name. By default, the community name is set to `XIV` (not public). To confirm or change the SNMP community name, click **System** → **Settings**, and then open the SNMP window that is shown in Figure 7-64.

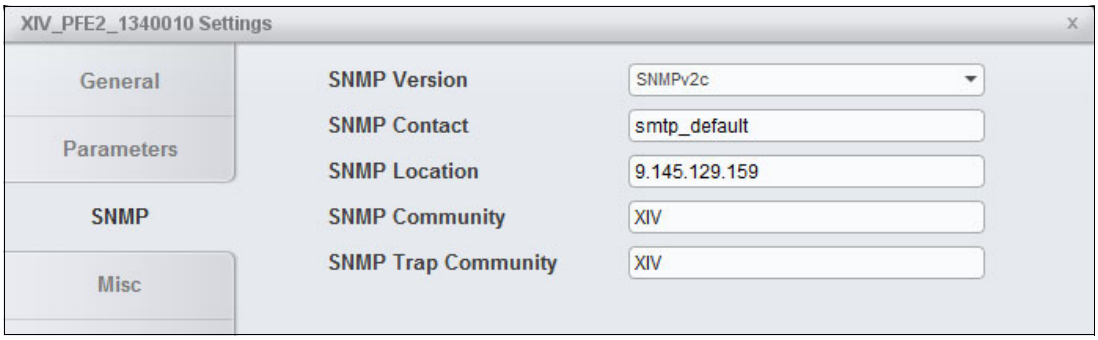


Figure 7-64 Set or show the XIV Storage System SNMP community name

You can also set or show SNMP information, including the community name, by running **`config_get`** and **`config_set`**, as shown in Example 7-42 on page 387.

Example 7-42 XIV Storage System SNMP commands

```
XIV_1312611>>config_get
Name                                    Value
dns_primary
dns_secondary
system_name
snmp_location
snmp_contact                            smtp_default
snmp_community                          XIV
snmp_trap_community                     XIV
snmp_type                               V2C
snmpv3_user
snmpv3_encryption_type                  AES
snmpv3_encryption_passphrase            ****
snmpv3_authentication_type              SHA
snmpv3_authentication_passphrase        ****
system_id                               40010
machine_type                            2810
machine_model                           214
machine_serial_number                   1340010
email_sender_address
email_reply_to_address
email_subject_format                    {severity}: {description}
internal_email_subject_format           {machine_type}-{machine_model}:....
iscsi_name                              iqn.2005-10.com.xivstorage:040010
maximal_snapshot_deletion_priority      4
timezone                                0
fc_proof                                yes
ntp_server
ups_control                             yes
support_center_port_type                Management
isns_server                             ?
ipv6_state                              enabled
ipsec_state                             disabled
ipsec_track_tunnels                     no
impending_power_loss_detection_method   UPS

XIV_1312611>>config_set name=snmp_location value=IBM_Mainz
Command executed successfully

XIV_1312611>>config_get name=snmp_location
Name            Value
snmp_location   IBM_Mainz
```

7.6.3 Using SNMP get or walk commands with open source software

You can test SNMP **get** and **walk** commands to an XIV Storage System using the open source software package `net-snmp` on a Windows workstation. This package provides a tool to compile the MIB and issue SNMP commands. You can download `net-snmp` from the Source Forge web page:

http://sourceforge.net/projects/net-snmp/files/net-snmp/

Consideration: Rtrieve the MIB contents from an XIV Storage System by running **mib_get**.

Force a compilation of all MIBs to compile the XIV Storage System MIB by running the following command:

```
C:\usr\bin>snmptranslate -Dparse-mibs
```

In the output that you get, look for messages such as the ones shown in Example 7-43. The module numbers might be different, depending on how many MIBs exist in that folder.

Example 7-43 MIB translation messages

```
XIV-MIB is in C:/usr/share/snmp/mibs/XIV-MIB-10.2.4.txt
Module 72 XIV-MIB is in C:/usr/share/snmp/mibs/xiv-10.2.x-mib.txt
Checking file: C:/usr/share/snmp/mibs/XIV-MIB-10.2.4.txt...
XIV-MIB is now in C:/usr/share/snmp/mibs/XIV-MIB-10.2.4.txt
```

Now you are ready to use an SNMP **walk** command starting with the `xiv` OID. The only thing that you need to change in this line is the management IP address of the XIV Storage System:

```
C:\usr\bin> snmpwalk -v 2c -c XIV -m XIV-MIB 10.10.1.10 xiv
```

In this command, the parameters have the following meanings:

-v 2c	Forces SNMP Version 2c.
-c XIV	The community name. It is case-sensitive.
-m XIV-MIB	Forces the use of the XIV Storage System MIB.
10.10.1.10	An XIV Storage System management module IP address.
xiv	The root of the XIV Storage System MIB.

In Example 7-44, you see some typical output. If your XIV Storage System is using Version 10.2.2 or earlier, you also get a list of XIV Storage System events. If that is the case, do not leave the **snmpwalk** command running; press Ctrl-c to stop it. XIV Storage System code Versions 10.2.4 and higher do not list XIV Storage System events through the **snmpwalk** command.

Example 7-44 The output of the snmpwalk command

```
C:\usr\bin>snmpwalk -v 2c -c XIV -m XIV-MIB 10.10.1.10 xiv
XIV-MIB::xivMachineStatus = STRING: "Full Redundancy"
XIV-MIB::xivFailedDisks = INTEGER: 0
XIV-MIB::xivUtilizationSoft = Gauge32: 69
XIV-MIB::xivUtilizationHard = Gauge32: 90
XIV-MIB::xivFreeSpaceSoft = INTEGER: 49770
XIV-MIB::xivFreeSpaceHard = INTEGER: 7954
XIV-MIB::xivIfIOPS.1004 = Gauge32: 6519
XIV-MIB::xivIfIOPS.1005 = Gauge32: 6773
XIV-MIB::xivIfIOPS.1006 = Gauge32: 6515
XIV-MIB::xivIfIOPS.1007 = Gauge32: 6557
XIV-MIB::xivIfIOPS.1008 = Gauge32: 6517
XIV-MIB::xivIfIOPS.1009 = Gauge32: 6575
XIV-MIB::xivIfStatus.1004 = STRING: "OK"
XIV-MIB::xivIfStatus.1005 = STRING: "OK"
XIV-MIB::xivIfStatus.1006 = STRING: "OK"
XIV-MIB::xivIfStatus.1007 = STRING: "OK"
```

```
XIV-MIB::xivIfStatus.1008 = STRING: "OK"
XIV-MIB::xivIfStatus.1009 = STRING: "OK"
```

The following fields are shown in Example 7-44 on page 388:

`xivMachineStatus`	Shows if a disk rebuild or redistribution is occurring.
`xivFailedDisks`	The number of failed disks in the XIV Storage System.
`xivUtilizationSoft`	The percentage of total soft space that is allocated to pools.
`xivUtilizationHard`	The percentage of total hard space that is allocated to pools.
`xivFreeSpaceSoft`	The amount of soft space that is deallocated in GB.
`xivFreeSpaceHard`	The amount of hard space that is deallocated in GB.
`xivIfIOPS.1004`	The number of IOPS being executed by module 4 at that moment.
`xivIfIOPS.1005`	The number of IOPS being executed by module 5 at that moment.
`xivIfIOPS.1006`	The number of IOPS being executed by module 6 at that moment.
`xivIfIOPS.1007`	The number of IOPS being executed by module 7 at that moment.
`xivIfIOPS.1008`	The number of IOPS being executed by module 8 at that moment.
`xivIfIOPS.1009`	The number of IOPS being executed by module 9 at that moment.
`xivIfStatus.1004`	The status of module 4 at that moment.
`xivIfStatus.1005`	The status of module 5 at that moment.
`xivIfStatus.1006`	The status of module 6 at that moment.
`xivIfStatus.1007`	The status of module 7 at that moment.
`xivIfStatus.1008`	The status of module 8 at that moment.
`xivIfStatus.1009`	The status of module 9 at that moment.

Tip: In the output shown in Example 7-44, you can see the `xivFreeSpaceSoft` output and `xivFreeSpaceHard` output. This information is only useful if you want to confirm how much space is not allocated to a pool. If you have already allocated all usable hard and soft space to your pools, this command confirms that there is no free space available outside your existing pools. There might be significant free space within your pools.

If you want to cut the output back to a single field (OID), use the syntax shown in Example 7-45. You can also run the **`snmpget`** command to get the same output.

Example 7-45 SNMP walks with a single OID

```
C:\usr\bin>snmpwalk -v 2c -c XIV -m XIV-MIB 10.10.1.10 xivFailedDisks
XIV-MIB::xivFailedDisks = INTEGER: 0

C:\usr\bin>snmpwalk -v 2c -c XIV -m XIV-MIB 10.10.1.10 xivFreeSpaceSoft
XIV-MIB::xivFreeSpaceSoft = INTEGER: 17111
```

If you start the SNMP **`walk`** without a start point, you get a more interesting output. You see the Linux version being run by the XIV Storage System modules and the uptime of the module you are probing, as shown in Example 7-46.

Example 7-46 Using the snmpwalk command without an OID

```
C:\usr\share\snmp\mibs>snmpwalk -v 2c -c XIV -m XIV-MIB 10.10.1.10
SNMPv2-SMI::mib-2.1.1.0 = STRING: "Linux nextra-WallStCO-module-4
2.6.16.46-268-xiv-220-x86_64-ixss10.2.4 #1 SMP Tue Nov 16 00:43:46 UTC 2010
x86_64"
SNMPv2-SMI::mib-2.1.2.0 = OID: SNMPv2-SMI::enterprises.8072.3.2.10
SNMPv2-SMI::mib-2.1.3.0 = Timeticks: (235647520) 27 days, 6:34:35.20
```

7.7 Using Tivoli Storage Productivity Center

Starting with Version 10.2.2 of the XIV Storage System Software, integration with Tivoli Storage Productivity Center v4.2 or higher also allows XIV Storage System performance monitoring from Tivoli Storage Productivity Center.

Tivoli Storage Productivity Center is an integrated suite for managing storage systems, capacity, storage networks, and replication. For information about Tivoli Storage Productivity Center, see *IIBM Tivoli Storage Productivity Center V5.2 Release Guide*, SG24-8204.

7.7.1 Setting up and discovering XIV systems in Tivoli Storage Productivity Center

If you are still using Tivoli Storage Productivity Center v4.1.*x*, Tivoli Storage Productivity Center manages and monitors the XIV Storage System through its CIM agent (embedded in the XIV Storage System code). You need to upgrade to Tivoli Storage Productivity Center v4.2 or later so that you can use the new data collection process and the Native API to interface with the XIV Storage System.

To add an XIV Storage System to your Tivoli Storage Productivity Center configuration, complete the following steps:

1. Create an XIV Storage System user and password that will be used by Tivoli Storage Productivity Center.
2. Note the IBM XIV Storage System management IP address.
3. Start the Tivoli Storage Productivity Center configure devices wizard to discover and probe the XIV Storage System.
4. Create a performance monitor.

XIV Storage System user creation

Create a normal XIV Storage System user and password on the XIV Storage System by using the XIV Storage Management GUI, as shown in Figure 7-65. This user and password is used in Tivoli Storage Productivity Center.

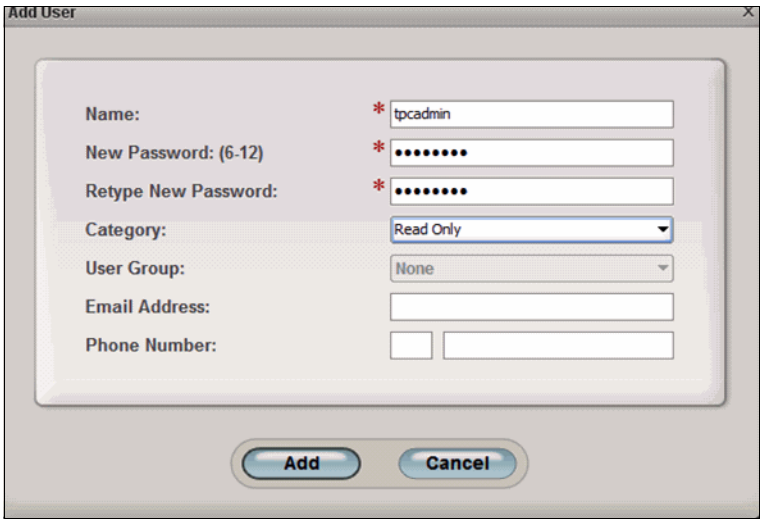


Figure 7-65 User creation

XIV Storage System management IP addresses

To add an XIV Storage System into Tivoli Storage Productivity Center, you must know the management IP addresses of the XIV Storage System modules 4, 5, and 6. Figure 7-66 shows how to view the management IP addresses. Right-click a particular XIV Storage System in the All Systems view from the XIV Storage Management GUI and select **Modify IP Addresses**.

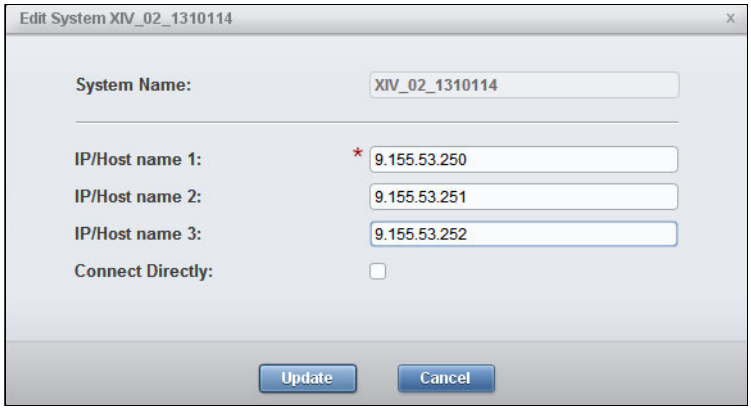


Figure 7-66 Showing the management IP addresses

Discovery phase

To add the XIV Storage System to Tivoli Storage Productivity Center, start the Configure Devices wizard. Click **IBM Tivoli Storage Productivity Center** → **Configure Devices**, as shown in Figure 7-67. Proceed with the steps to add, discover, and probe an XIV Storage System in Tivoli Storage Productivity Center.

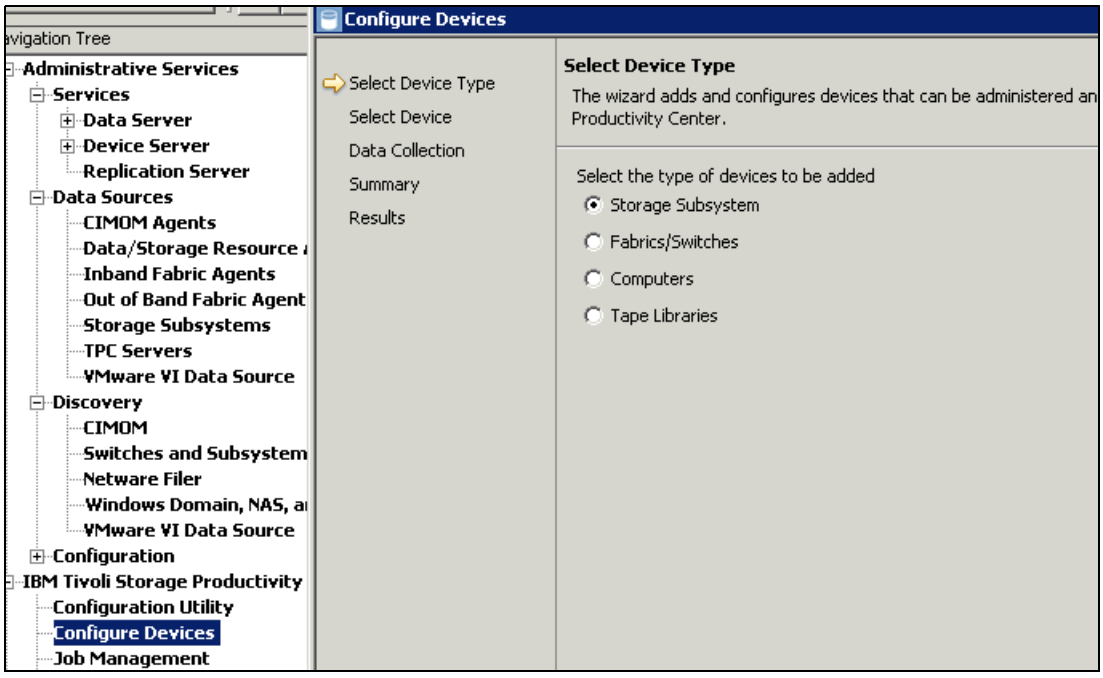


Figure 7-67 Configure Devices wizard

The discovery usually takes a few minutes and can be run on a schedule. How often you run a discovery depends on the dynamic of your environment. It must be run to detect a new subsystem and to perform basic health checks of all CIMOMs and other storage subsystems.

Tivoli Storage Productivity Center configuration wizard

To use the Tivoli Storage Productivity Center wizard, complete the following steps:

1. Click **IBM Tivoli Storage Productivity Center** → **Configure Devices** to start the wizard.
2. Click **Storage Subsystem** and click **Next**.
3. Click **Add and configure new storage subsystem** and click **Next**.
4. Click **IBM XIV** as the Device Type and enter the required information:
   - IP Address: For the XIV Storage System, this address corresponds to the three management IP address of the XIV Storage System. Start with the first IP address.
   - Username: The XIV Storage System user that is created for Tivoli Storage Productivity Center.
   - Password: The password of the XIV Storage System.
5. Click **Add** and complete step 4 for the other IP addresses, as shown in Figure 7-68 on page 392.

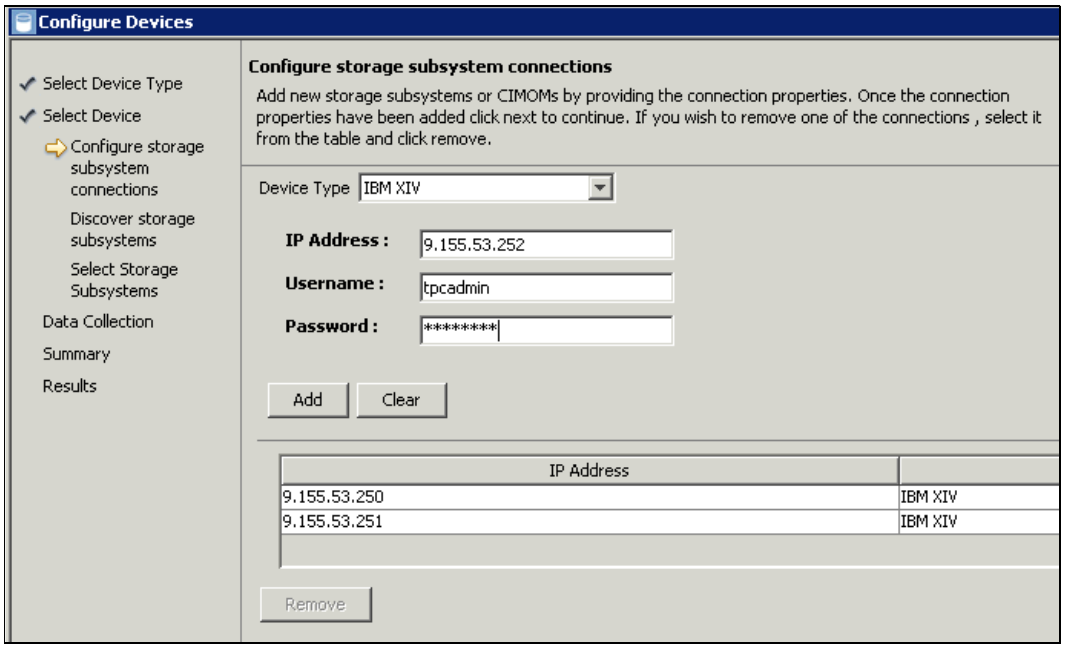


Figure 7-68 Configure storage subsystem connections

6. After all three IP addresses are added, click **Next** and the wizard automatically performs a discovery.

7. Click **Next** again and the discovered XIV Storage System is shown as unconfigured storage, as shown in Figure 7-69.

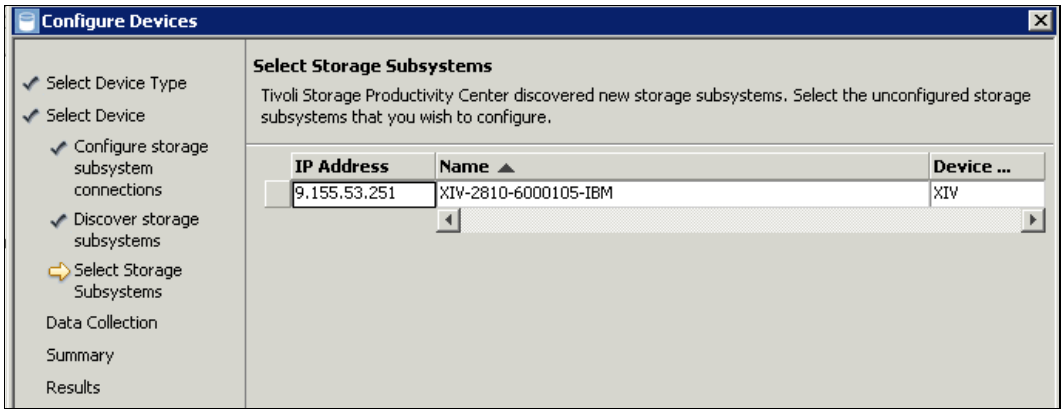


Figure 7-69 New XIV Storage System discovered in Tivoli Storage Productivity Center

Probing phase

The newly added XIV Storage System must be probed for Tivoli Storage Productivity Center to collect information. Probes use agents to collect statistics, including data about drives, pools, and volumes. The results of the probe jobs are stored in the repository and are used in Tivoli Storage Productivity Center to supply the data necessary for generating several reports, including Asset, Capacity, and Storage Subsystem reports.

To configure the probe for XIV Storage System, continue the following steps in the wizard:

1. Select the XIV Storage System and click **Next**. Now you have various options to specify the probe details, as shown in Figure 7-70 on page 393.

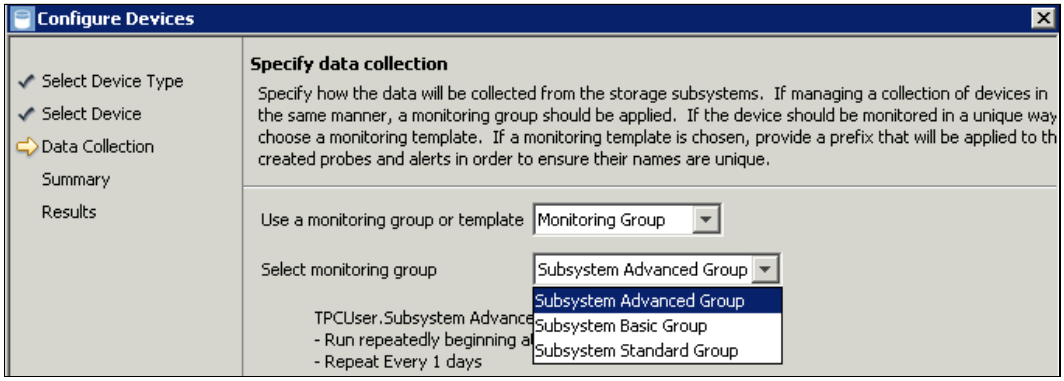

Configure Devices
Select Device Type
Select Device
Data Collection
Summary
Results
Specify data collection
Specify how the data will be collected from the storage subsystems. If managing a collection of devices in the same manner, a monitoring group should be applied. If the device should be monitored in a unique way choose a monitoring template. If a monitoring template is chosen, provide a prefix that will be applied to th created probes and alerts in order to ensure their names are unique.
Use a monitoring group or template
Monitoring Group
Select monitoring group
Subsystem Advanced Group
Subsystem Advanced Group
Subsystem Basic Group
Subsystem Standard Group
TPCUser.Subsystem Advance
- Run repeatedly beginning a
- Repeat Every 1 days

Figure 7-70 Specifying probe details

Specify the probe details:

- Use a monitoring group or template: Select the predefined **Monitoring Group** or create your own new group by selecting **Monitoring Template**.
- Select monitoring group: Specify **Subsystem Advanced Group**, **Subsystem Basic Group**, or **Subsystem Standard Group**.

> **Tip:** Configure individual probes for every XIV Storage System system, but set them to run at various times.

After clicking **Next**, you get an overview of your configured probe. Figure 7-71 shows the details of the Subsystem Standard Group, including the scheduling of the probe and the alerts.

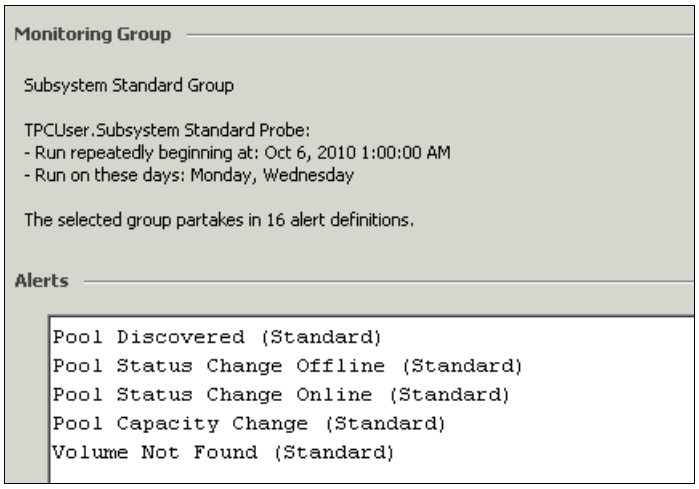

Monitoring Group
Subsystem Standard Group
TPCUser.Subsystem Standard Probe:
- Run repeatedly beginning at: Oct 6, 2010 1:00:00 AM
- Run on these days: Monday, Wednesday
The selected group partakes in 16 alert definitions.
Alerts
Pool Discovered (Standard)
Pool Status Change Offline (Standard)
Pool Status Change Online (Standard)
Pool Capacity Change (Standard)
Volume Not Found (Standard)

Figure 7-71 Probe summary

2. Confirm the summary by clicking **Next**.
3. Click **Finish** and a background probe is started automatically.
4. Click **IBM Tivoli Storage Productivity Center** → **Configuration Utility** → **Disk Manager** to check the status of your XIV Storage System. You can also use the Disk Manager to run a new discovery to create a probe or to modify the alerts (Figure 7-72 on page 394).

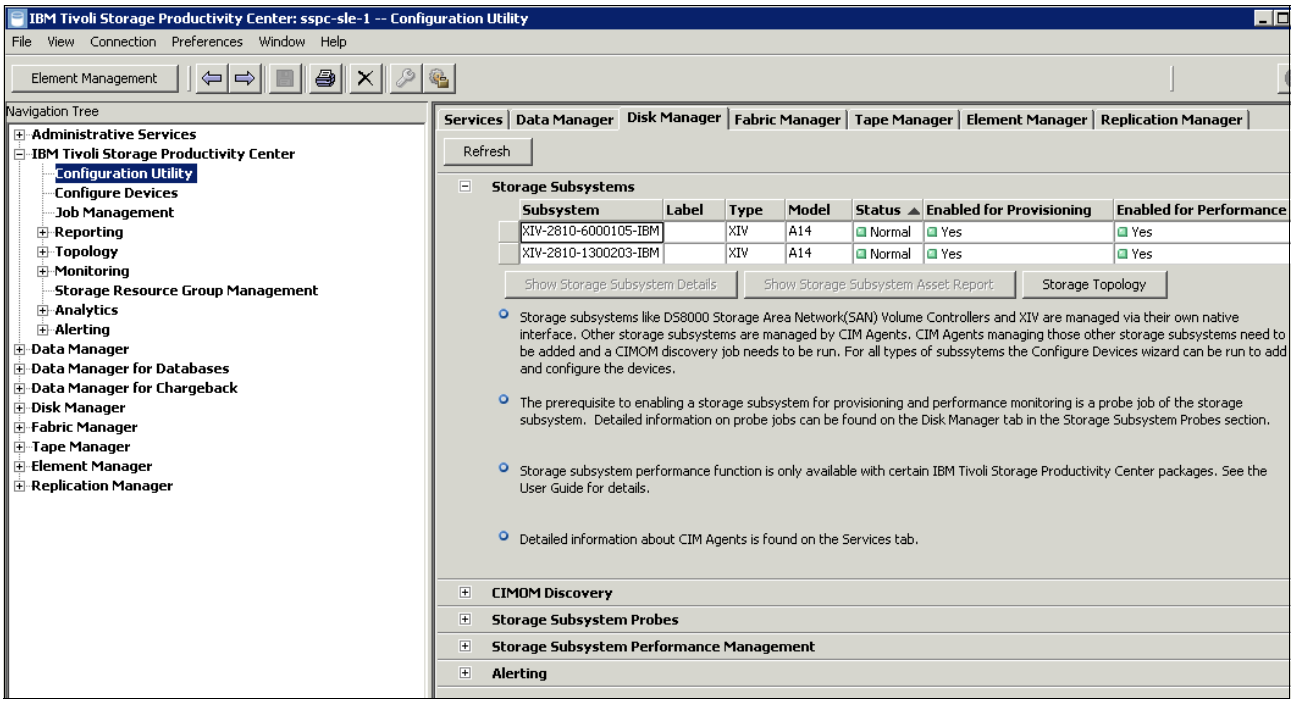

IBM Tivoli Storage Productivity Center: sspc-sle-1 -- Configuration Utility
File View Connection Preferences Window Help
Element Management
Navigation Tree
Administrative Services
IBM Tivoli Storage Productivity Center
Configuration Utility
Configure Devices
Job Management
Reporting
Topology
Monitoring
Storage Resource Group Management
Analytics
Alerting
Data Manager
Data Manager for Databases
Data Manager for Chargeback
Disk Manager
Fabric Manager
Tape Manager
Element Manager
Replication Manager
Services Data Manager Disk Manager Fabric Manager Tape Manager Element Manager Replication Manager
Refresh
Storage Subsystems
Subsystem Label Type Model Status Enabled for Provisioning Enabled for Performance
XIV-2810-6000105-IBM XIV A14 Normal Yes Yes
XIV-2810-1300203-IBM XIV A14 Normal Yes Yes
Show Storage Subsystem Details
Show Storage Subsystem Asset Report
Storage Topology
Storage subsystems like DS8000 Storage Area Network(SAN) Volume Controllers and XIV are managed via their own native interface. Other storage subsystems are managed by CIM Agents. CIM Agents managing those other storage subsystems need to be added and a CIMOM discovery job needs to be run. For all types of subssytems the Configure Devices wizard can be run to add and configure the devices.
The prerequisite to enabling a storage subsystem for provisioning and performance monitoring is a probe job of the storage subsystem. Detailed information on probe jobs can be found on the Disk Manager tab in the Storage Subsystem Probes section.
Storage subsystem performance function is only available with certain IBM Tivoli Storage Productivity Center packages. See the User Guide for details.
Detailed information about CIM Agents is found on the Services tab.
CIMOM Discovery
Storage Subsystem Probes
Storage Subsystem Performance Management
Alerting

Figure 7-72 Disk Manager tab

If you want to modify the probe you created, click **IBM Tivoli Storage Productivity Center** → **Monitoring** → **Probes** and choose the correct group (in this example, **Subsystem Standard Probe**). Click **When to Run** to adjust the schedule or **Alert** if you want to modify the alerts, as shown in Figure 7-73.

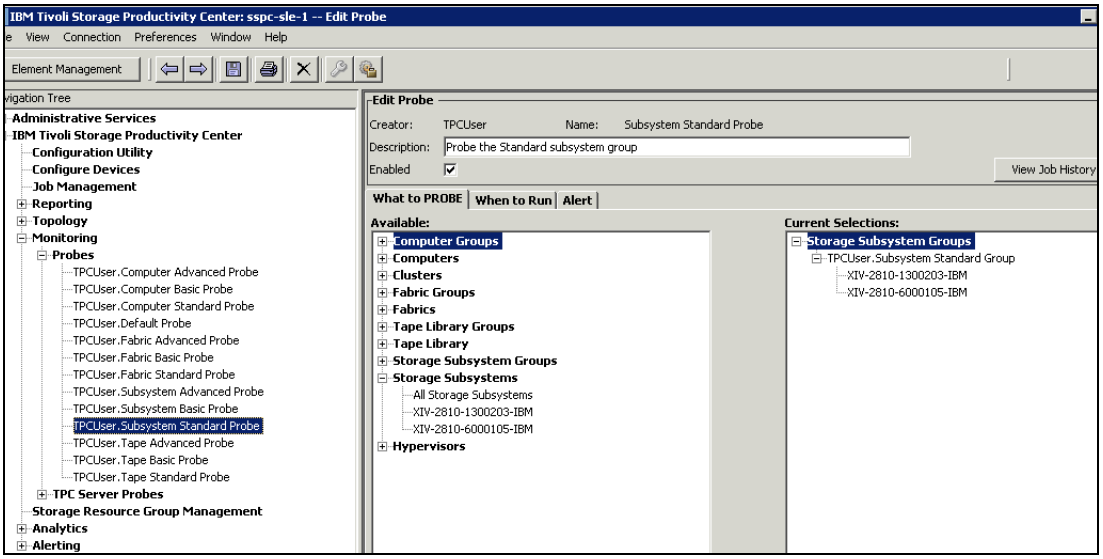


Figure 7-73 Verifying the probe details

Performance monitoring

XIV Storage System Software v10.2.2 or later can use the performance monitoring feature in Tivoli Storage Productivity Center v4.2. You can set up this feature with the Disk Manager.

Complete the following steps:

1. Click **IBM Tivoli Storage Productivity Center** → **Configuration Utility** → **Disk Manager** tab → **Subsystem Performance Monitors** → **Create Subsystem Performance Monitor** to create a performance monitor, as shown in Figure 7-74.

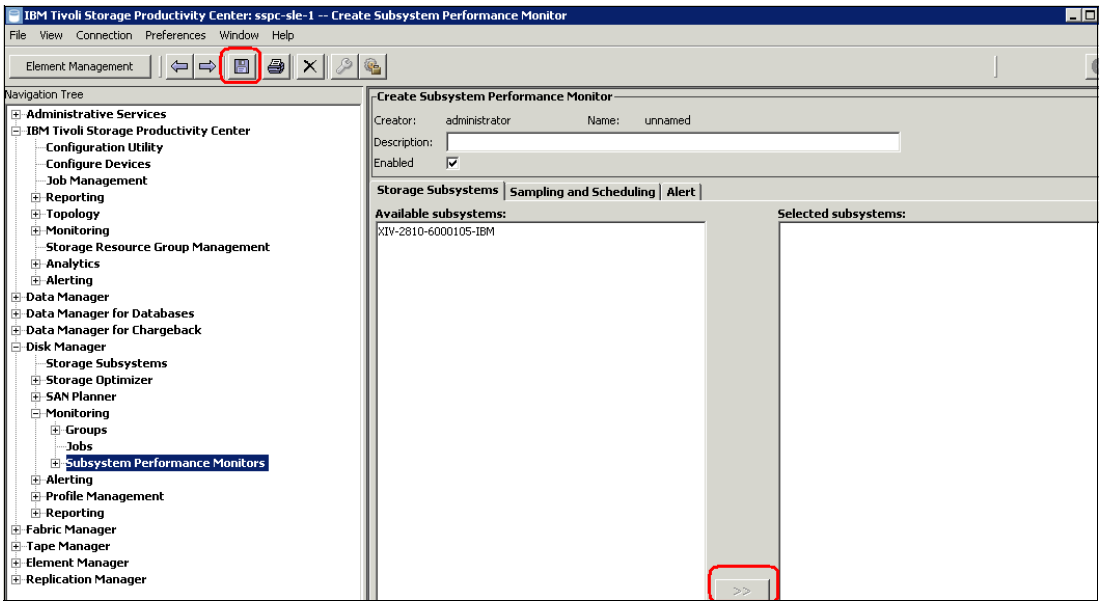


Figure 7-74 Subsystem Performance Monitors panel

Optional: You can also click **Disk Manager** → **Monitoring** → **Subsystem Performance Monitors** to open the Subsystem Performance Monitors view.

2. Afterward, select the XIV Storage System and click the double arrow icon to move the selected system from the Available subsystems into the Selected subsystems section.
3. Click **Save** and specify the subsystem performance monitor name. Afterward, the subsystem performance monitor is automatically submitted.
4. If you want to review or change the sampling and scheduling, select the performance monitor, click **Disk Manager** → **Monitoring** → **Subsystem Performance Monitors** and click **Sampling and Scheduling**, as shown in Figure 7-75.

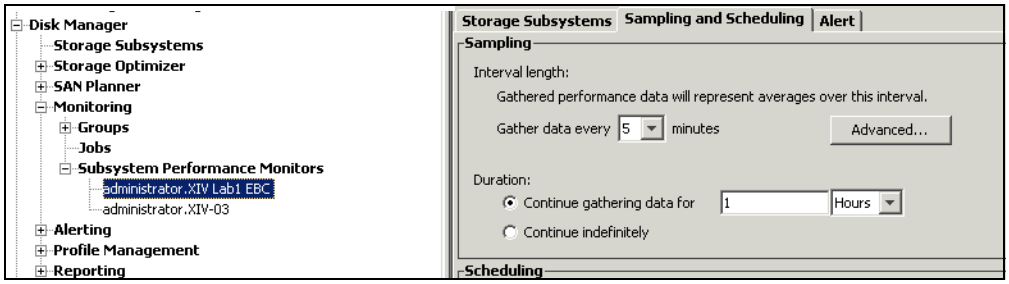


Figure 7-75 Performance Monitor details

Configuration information and reporting

Figure 7-76 shows a list of various XIV Storage System subsystems as reported in Tivoli Storage Productivity Center.

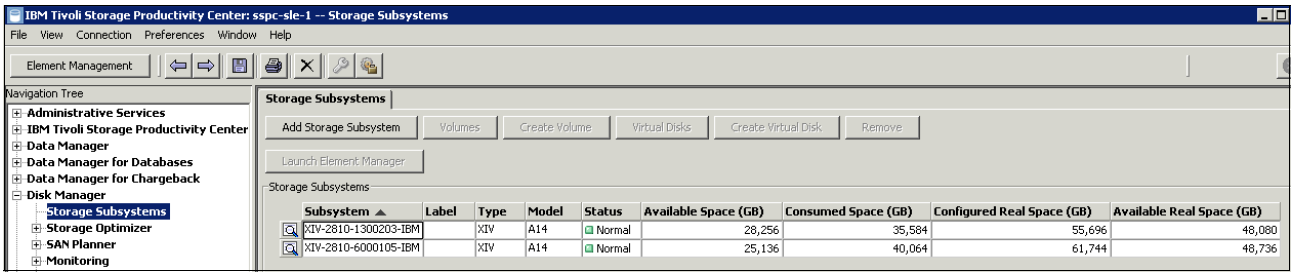


Figure 7-76 List storage subsystems

7.7.2 XIV Storage System Tivoli Storage Productivity Center reports

Tivoli Storage Productivity Center v4.2 includes basic capacity and asset information in tabular reports and in the topology viewer. In addition, logical unit number (LUN) correlation information and performance data is available.

Tivoli Storage Productivity Center probes collect the following information from the IBM XIV systems:

- Storage pools
- Volumes
- Disks
- Ports
- Host definitions, LUN mapping, and masking information

Capacity: Space is calculated differently in the XIV Storage Management GUI and the XCLI than in Tivoli Storage Productivity Center:

- XIV defines space this way:
  - 1 GB as 10^9 = 1,000,000,000 bytes
  - 1 TB as 10^{12} = 1,000,000,000,000 bytes
- Tivoli Storage Productivity Center defines space this way:
  - 1 GiB as 2^{30} = 1,073,741,824 bytes
  - 1 TiB as 2^{40} = 1,099,511,627,776 bytes

These different definitions are why capacity information might appear to differ (and seem wrong) when comparing the XIV Storage Management GUI with the Tivoli Storage Productivity Center GUI, when in fact it is the same size.

Because the XIV Storage System also provides thin provisioning, additional columns for the thin provisioning storage pools were introduced to the Tivoli Storage Productivity Center GUI.

The Tivoli Storage Productivity Center *configured space* is equivalent to the XIV Storage System soft capacity, and Tivoli Storage Productivity Center *real space* is equivalent to the XIV Storage System hard space.

Additional Configured Capacity Limit and Remaining Configured Capacity columns were introduced to report on the hard capacity of a subsystem. The pre-existing Consumed Space and Available Space columns now report on the soft capacity of a subsystem in the following reports:

- Storage Subsystem list (Click **Disk Manager** → **Storage Subsystems** to view this report.)
- Storage Subsystem Details (Click **Disk Manager** → **Storage Subsystems** and select a storage subsystem to view this report.)
- Storage Subsystem Details (Click **Data Manager** → **Reporting** → **Asset** → **By Storage Subsystem** to view storage subsystem details.)
- Storage Subsystem Reports can be created with user-specified columns. See the following Tivoli Storage Productivity Center GUI menu items:
  - **Data Manager** → **Reporting** → **Asset** → **System-wide** → **Storage Subsystems**
  - **Data Manager** → **Reporting** → **TPC-wide Storage Space** → **Disk Space** → **By Storage Subsystem**

In Figure 7-77, the properties of an XIV Storage System are shown in the Tivoli Storage Productivity Center GUI Storage Subsystem Details report.

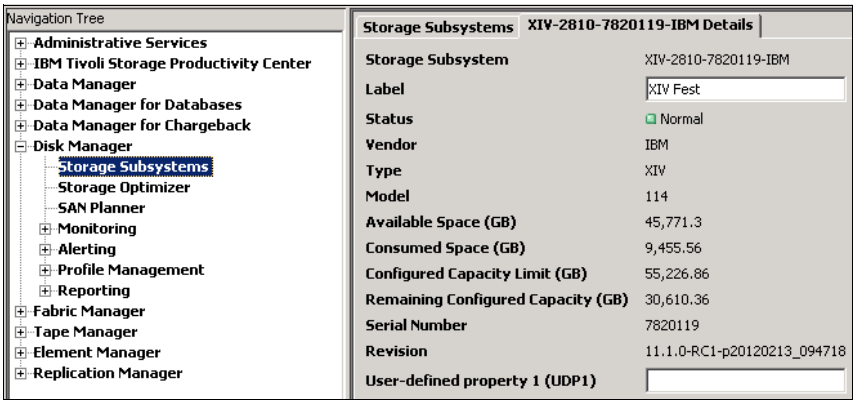


Figure 7-77 XIV Storage System details in the Tivoli Storage Productivity Center GUI

The fields in Figure 7-77 show the following information:

Available Space (GB) The combined unused soft space inside all defined pools (in GiB).

Consumed Space (GB) The combined size of all defined volumes and snapshot reserves in all pools (in GiB).

Configured Capacity Limit (GB) The total of the hard size of all defined pools (in GiB). All of the hard space inside all of the pools.

Remaining Configured Capacity (GB) The total of the remaining hard space (space not used by volumes or snapshots) inside all defined pools (in GiB).

Figure 7-78 on page 398 shows the details for an XIV Storage System as shown in Tivoli Storage Productivity Center. Do not use the fields that show the Disk Space, Available Disk Space, Physical Disk Space, Formatted Space, and Formatted Space with No Volumes. Although values are shown for some of these fields, the methods used to calculate them do not apply well to the XIV Storage System. They are used for other storage products.

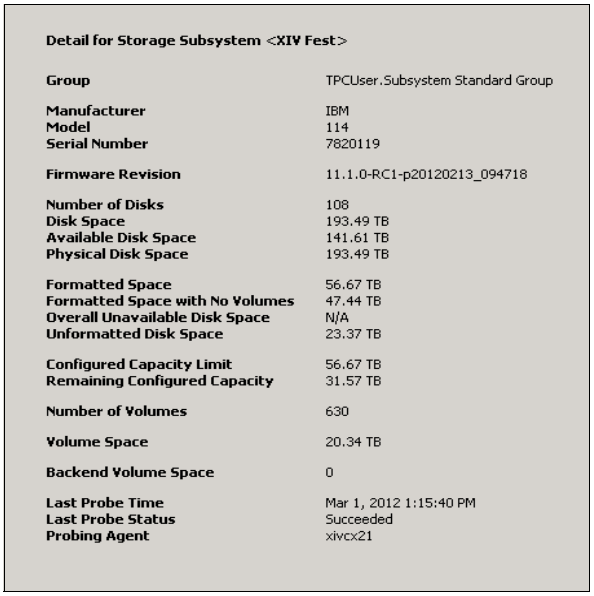


Figure 7-78 XIV Storage System details shown in Tivoli Storage Productivity Center

The remaining fields in Figure 7-78 show the following information:

Unformatted Disk Space The total available hard space that is left for pool creation

Configured Capacity Limit (GB) The combined hard size of all defined pools in TiB (which is GiB divided by 1024)

Remaining Configured Capacity (GB) The remaining hard space (space not used by volumes or snapshots) in all defined pools in TiB (which is GiB divided by 1024)

Tip: In Figure 7-78 on page 398, you can see the Last Probe Time, showing when Tivoli Storage Productivity Center last communicated with the XIV Storage System. Suppose that the information showing in Tivoli Storage Productivity Center differs from the information shown in the XIV Storage Management GUI or XCLI, even after converting from binary GiB (or TiB) to decimal GB (or TB). If this situation occurs, you might need to run a fresh probe in Tivoli Storage Productivity Center to update the information being shown in Tivoli Storage Productivity Center.

Determining the available and total hard space using Tivoli Storage Productivity Center

A common task is to analyze how much space is available for configuration. Focusing on hard space, there are three useful measures shown in Figure 7-78:

- *Unformatted Disk Space* shows how much hard space is available to create new pools.
- If you are planning to convert your pools to thin provisioning pools, you can use *Remaining Configured Capacity* to determine how much hard space remains in the existing pools.
- To determine total hard space in your XIV Storage System, you can sum together *Unformatted Disk Space* and *Configured Capacity Limit.*

  Because these two values are in binary TB (TiB), the resulting value does not match the decimal GB value shown for hard space in the XIV Storage Management GUI. A simple conversion is to multiply the TiB value shown in Tivoli Storage Productivity Center by 1.0995 to generate a decimal TB value close to the one shown in the XIV Storage Management GUI (variations occur because Tivoli Storage Productivity Center rounds the values down).

Table 7-3 shows the usable hard space in binary TiB that is reported by Tivoli Storage Productivity Center. The value varies based on the model of XIV Storage System, the size of the drives being used, and the number of modules that are physically installed. Variation by 0.01 TB might occur due to rounding.

Table 7-3 XIV total hard space in binary TiB as reported by Tivoli Storage Productivity Center

Modules	XIV Gen3 with 2 TB	XIV Gen3 with 3 TB
6	50.66	76.53
9	80.03	120.83
10	93.34	140.88
11	101.44	153.11
12	114.56	172.87
13	122.73	185.19

Modules	XIV Gen3 with 2 TB	XIV Gen3 with 3 TB
14	135.83	204.94
15	146.72	221.36

Storage pools

Tivoli Storage Productivity Center can report on XIV Storage System storage pools. To assist with this action, new Configured Real Space and Available Real Space columns, reporting on the hard capacity of a storage pool, were added to the Storage Pool Details report. This report can be accessed by clicking **Data Manager** → **Reporting** → **Asset** → **By Storage Subsystem** → ***<Subsystem Name>*** → **Storage Pools**.

To demonstrate how to interpret the values shown in Tivoli Storage Productivity Center, see Figure 7-79, where an XIV Storage System pool called *ITSO* was created. The hard size of the pool is 5016 GB and the soft size of the pool is 10015 GB (shown in the XIV Storage Management GUI as 10 TB). There is one volume in the pool, sized at 1099 GB. In that volume is 533 GB of actual data.

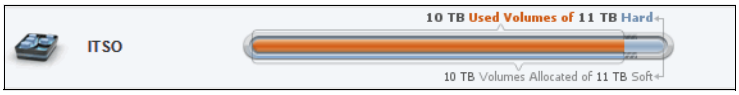


Figure 7-79 ITSO pool shown in XIV Storage Management GUI

In Figure 7-80 on page 400, the details of the ITSO pool are shown in the Tivoli Storage Productivity Center GUI.

Detail for Storage Pool	ITSO
Storage Subsystem	XIV Fest
Type	RAID 10
Status	OK
Storage Pool Space	10.27 TB
Available Storage Pool Space	1.03 TB
Configured Capacity Limit	10.27 TB
Remaining Configured Capacity	961.58 GB
Is Space Efficient	true
Is pool encryptable	false
Is pool encrypted	false
Solid State	Non-Solid State
Easy Tier	N/A
Easy Tier Status	N/A
Number of Disks	N/A
Number of Volumes	10
Surfaced Volume Space	9.23 TB
Un-surfaced Volume Space	0
Backend Volume Space	0

Figure 7-80 XIV Storage System ITSO storage pool details in Tivoli Storage Productivity Center GUI

The values shown in Figure 7-79 on page 400 and Figure 7-80 correspond in the following way:

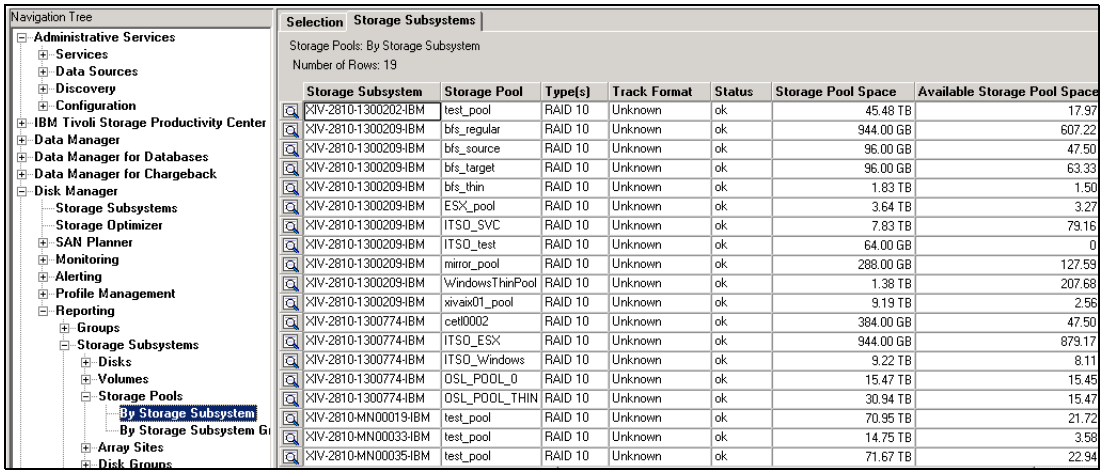
Storage Pool Space The soft size of the pool (10015 GB equals 9.11 TiB)

Available Storage Pool Space The unused soft space in the pool (in binary TiB)

Configured Capacity Limit The hard size of the pool (5016 GB equals 4.56 TiB)

Remaining Configured Capacity The unused hard size in the pool (in binary TiB)

Figure 7-81 shows how Tivoli Storage Productivity Center can report on XIV Storage System storage pools.



Storage Subsystem	Storage Pool	Type(s)	Track Format	Status	Storage Pool Space	Available Storage Pool Space
XIV-2810-1300202-IBM	test_pool	RAID 10	Unknown	ok	45.48 TB	17.97
XIV-2810-1300209-IBM	bfs_regular	RAID 10	Unknown	ok	944.00 GB	607.22
XIV-2810-1300209-IBM	bfs_source	RAID 10	Unknown	ok	96.00 GB	47.50
XIV-2810-1300209-IBM	bfs_target	RAID 10	Unknown	ok	96.00 GB	63.33
XIV-2810-1300209-IBM	bfs_thin	RAID 10	Unknown	ok	1.83 TB	1.50
XIV-2810-1300209-IBM	ESX_pool	RAID 10	Unknown	ok	3.64 TB	3.27
XIV-2810-1300209-IBM	ITSO_SVC	RAID 10	Unknown	ok	7.83 TB	79.16
XIV-2810-1300209-IBM	ITSO_test	RAID 10	Unknown	ok	64.00 GB	0
XIV-2810-1300209-IBM	mirror_pool	RAID 10	Unknown	ok	288.00 GB	127.59
XIV-2810-1300209-IBM	WindowsThinPool	RAID 10	Unknown	ok	1.38 TB	207.68
XIV-2810-1300209-IBM	xivaix01_pool	RAID 10	Unknown	ok	9.19 TB	2.56
XIV-2810-1300774-IBM	cetl0002	RAID 10	Unknown	ok	384.00 GB	47.50
XIV-2810-1300774-IBM	ITSO_ESX	RAID 10	Unknown	ok	944.00 GB	879.17
XIV-2810-1300774-IBM	ITSO_Windows	RAID 10	Unknown	ok	9.22 TB	8.11
XIV-2810-1300774-IBM	OSL_POOL_0	RAID 10	Unknown	ok	15.47 TB	15.45
XIV-2810-1300774-IBM	OSL_POOL_THIN	RAID 10	Unknown	ok	30.94 TB	15.47
XIV-2810-MN00019-IBM	test_pool	RAID 10	Unknown	ok	70.95 TB	21.72
XIV-2810-MN00033-IBM	test_pool	RAID 10	Unknown	ok	14.75 TB	3.58
XIV-2810-MN00035-IBM	test_pool	RAID 10	Unknown	ok	71.67 TB	22.94

Figure 7-81 XIV Storage System storage pools as seen by Tivoli Storage Productivity Center

Volumes

A *Volume Real Space* column was added to report on the hard capacity of a volume. The pre-existing *Volume Space* columns report on the soft capacity of a volume in these reports:

- Volume Details window under **Disk Manager** → **Storage Subsystems** → **Volumes**
- **Disk Manager** → **Reporting** → **Storage Subsystems** → **Volumes**
- **Disk Manager** → **Reporting** → **Storage Subsystems** → **Volume to HBA Assignment**
- Added **Backend Volume Real Space** for XIV Storage System volumes as back-end volumes under **Disk Manager** → **Reporting** → **Storage Subsystems** → **Volume to Backend Volume Assignment**
- Volume Details window under **Data Manager** → **Reporting** → **Asset** → **By Storage Subsystem** → ***<Subsystem Name>*** → **Volumes**
- **Data Manager** → **Reporting** → **Asset** → **System-wide** → **Volumes**

In Figure 7-82, a volume is shown in the ITSO pool. It is 1015 GB in size and contains 1015 GB of actual data.

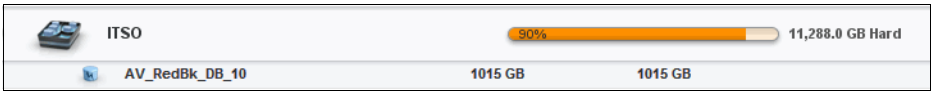


Figure 7-82 Volume properties shown in the XIV Storage Management GUI

In Figure 7-80 on page 400, the same volume is shown in the Tivoli Storage Productivity Center GUI. The volume space shows as 945.56 GB in Tivoli Storage Productivity Center because 1015 GB equals 945.56 GiB. The XIV Storage Management GUI rounds down the volume size. See Figure 7-83.

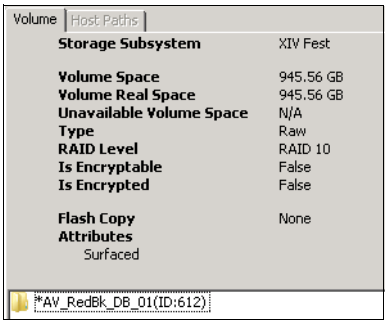


Figure 7-83 Volume properties shown in the Tivoli Storage Productivity Center GUI

Because of the nature of the XIV Storage System architecture and the fact that each volume is on all disks, various reports in the Tivoli Storage Productivity Center GUI do *not* necessarily provide meaningful information for IBM XIV systems. Correlation of disks and volumes, for example, under the **Data Manager** → **Reporting** → **Asset** → **By Storage Subsystem** → **Select a Storage Subsystem** → **Disks** branch, is not possible. Tivoli Storage Productivity Center does not report any volumes under the branch of a particular disk.

Also, because the XIV Storage System storage pools are used to group volumes but not disks, no disks are reported for a particular storage pool under that same reporting branch.

Finally, the following reports do not contain any information for XIV Storage System subsystems:

- **Disk Manager** → **Reporting** → **Storage Subsystems** → **Computer Views** → **By Computer** (Relate Computers to Disks)
- **Disk Manager** → **Reporting** → **Storage Subsystems** → **Computer Views** → **By Computer Group** (Relate Computers to Disks)
- **Disk Manager** → **Reporting** → **Storage Subsystems** → **Computer Views** → **By Filesystem/Logical Volume** (Relate Filesystems/Logical Volumes to Disks)
- **Disk Manager** → **Reporting** → **Storage Subsystems** → **Computer Views** → **By Filesystem Group** (Relate Filesystems/Logical Volumes to Disks)
- **Disk Manager** → **Reporting** → **Storage Subsystems** → **Storage Subsystem Views** → **Disks** (Relate Disks to Computers)

These queries, when combined with the Small Computer System Interface (SCSI) inquiry data that Tivoli Storage Productivity Center collects from the hosts, allow Tivoli Storage Productivity Center to correlate LUNs reported by the XIV Storage System to LUNs seen by the host systems.

Also, when the XIV Storage System is providing storage to the IBM System Storage SAN Volume Controller, Tivoli Storage Productivity Center can correlate LUNs reported by the XIV Storage System to SAN Volume Controller managed disks (MDisks).

Element Manager launch

If the XIV Storage Management GUI is installed and configured, Tivoli Storage Productivity Center element manager can be used to launch the XIV management software.

Tivoli Storage Productivity Center performance reports

The XIV Storage System Software v10.2.2 or later enables Tivoli Storage Productivity Center to create performance reports. The corresponding XIV Storage System performance monitors must be set up first, as described in 7.7.1, “Setting up and discovering XIV systems in Tivoli Storage Productivity Center” on page 390.

To create a performance report, click **IBM Tivoli Storage Productivity Center** → **Configuration Utility** → **Disk Manager** → **Storage Subsystem Performance Management** → **Show Subsystem Performance System Report**, as shown in Figure 7-84.

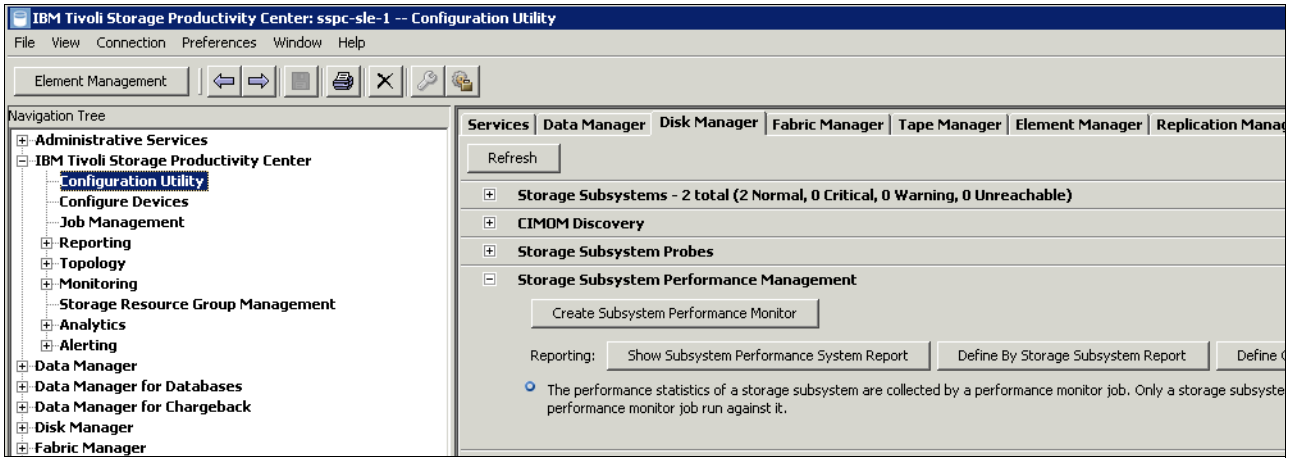

IBM Tivoli Storage Productivity Center: sspc-sle-1 -- Configuration Utility
File View Connection Preferences Window Help
Element Management
Navigation Tree
Administrative Services
IBM Tivoli Storage Productivity Center
Configuration Utility
Configure Devices
Job Management
Reporting
Topology
Monitoring
Storage Resource Group Management
Analytics
Alerting
Data Manager
Data Manager for Databases
Data Manager for Chargeback
Disk Manager
Fabric Manager
Services Data Manager Disk Manager Fabric Manager Tape Manager Element Manager Replication Manag
Refresh
Storage Subsystems - 2 total (2 Normal, 0 Critical, 0 Warning, 0 Unreachable)
CIMOM Discovery
Storage Subsystem Probes
Storage Subsystem Performance Management
Create Subsystem Performance Monitor
Reporting:
Show Subsystem Performance System Report
Define By Storage Subsystem Report
The performance statistics of a storage subsystem are collected by a performance monitor job. Only a storage subsyste
performance monitor job run against it.

Figure 7-84 Show performance report

Afterward, you can drill down into the details of any storage subsystem or even deeper into the single volume view and create Tivoli Storage Productivity Center reports or graphs, as shown in Figure 7-85 on page 403.

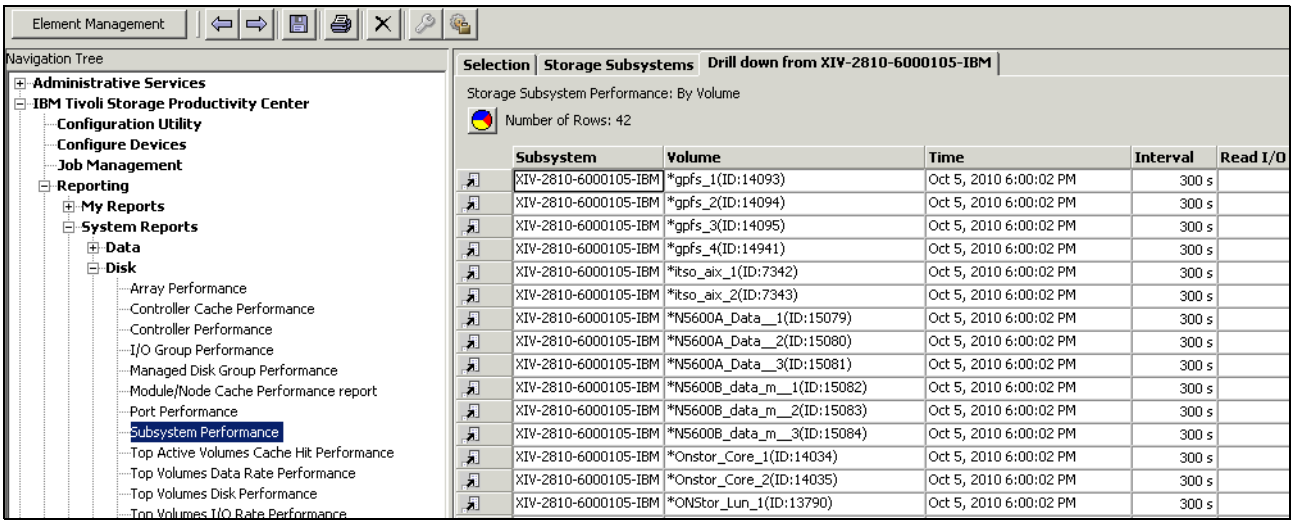

Element Management
Navigation Tree
Administrative Services
IBM Tivoli Storage Productivity Center
Configuration Utility
Configure Devices
Job Management
Reporting
My Reports
System Reports
Data
Disk
Array Performance
Controller Cache Performance
Controller Performance
I/O Group Performance
Managed Disk Group Performance
Module/Node Cache Performance report
Port Performance
Subsystem Performance
Top Active Volumes Cache Hit Performance
Top Volumes Data Rate Performance
Top Volumes Disk Performance
Top Volumes I/O Rate Performance
Selection Storage Subsystems Drill down from XIV-2810-6000105-IBM
Storage Subsystem Performance: By Volume
Number of Rows: 42
Subsystem Volume Time Interval Read I/O
XIV-2810-6000105-IBM *gpfs_1(ID:14093) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *gpfs_2(ID:14094) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *gpfs_3(ID:14095) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *gpfs_4(ID:14941) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *itso_aix_1(ID:7342) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *itso_aix_2(ID:7343) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *N5600A_Data__1(ID:15079) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *N5600A_Data__2(ID:15080) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *N5600A_Data__3(ID:15081) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *N5600B_data_m__1(ID:15082) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *N5600B_data_m__2(ID:15083) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *N5600B_data_m__3(ID:15084) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *Onstor_Core_1(ID:14034) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *Onstor_Core_2(ID:14035) Oct 5, 2010 6:00:02 PM 300 s
XIV-2810-6000105-IBM *ONStor_Lun_1(ID:13790) Oct 5, 2010 6:00:02 PM 300 s

Figure 7-85 Subsystem performance

Tivoli Storage Productivity Center volume creation

After the Configure Devices tool is used to add an XIV Storage System into Tivoli Storage Productivity Center, the NAPI introduced with Tivoli Storage Productivity Center v4.2 is used as the native interface between Tivoli Storage Productivity Center and the XIV Storage System. This interface also enables Tivoli Storage Productivity Center to create a volume on the XIV Storage System and to map it to a host by using the Volume Create wizard.

Choose the XIV Storage System for which you want to create a volume by clicking **Disk Manager** → **Storage Subsystem**, and then click **Create Volumes** to start the Volume Create wizard.

The Tivoli Storage Productivity Center XIV user ID must have the *storageadmin* authority to create volumes.

7.7.3 Tivoli Storage Productivity Center web-based GUI

Tivoli Storage Productivity Center can be accessed by using a web browser. This feature is part of the normal installation and creates a link in the All Programs menu named *Tivoli Storage Productivity Center web-based GUI*. This opens a browser window using a link, such as `https://ipaddress:9569/srm/Login.jsp`, where *ipaddress* is the address of the Tivoli Storage Productivity Center server.

This new GUI interface, available with Tivoli Storage Productivity Center 5.1, offers the now common look and feel already available with other IBM Storage System interfaces. Convenient links to related information make navigation quick and easy. The following details give you an overview of the features that are available in the new web-based GUI as it relates to the XIV Storage System.

The Home - Dashboard window gives a general view of the current configured resources and their status. From this window, you can select various options for displaying more information, as depicted in Figure 7-86.

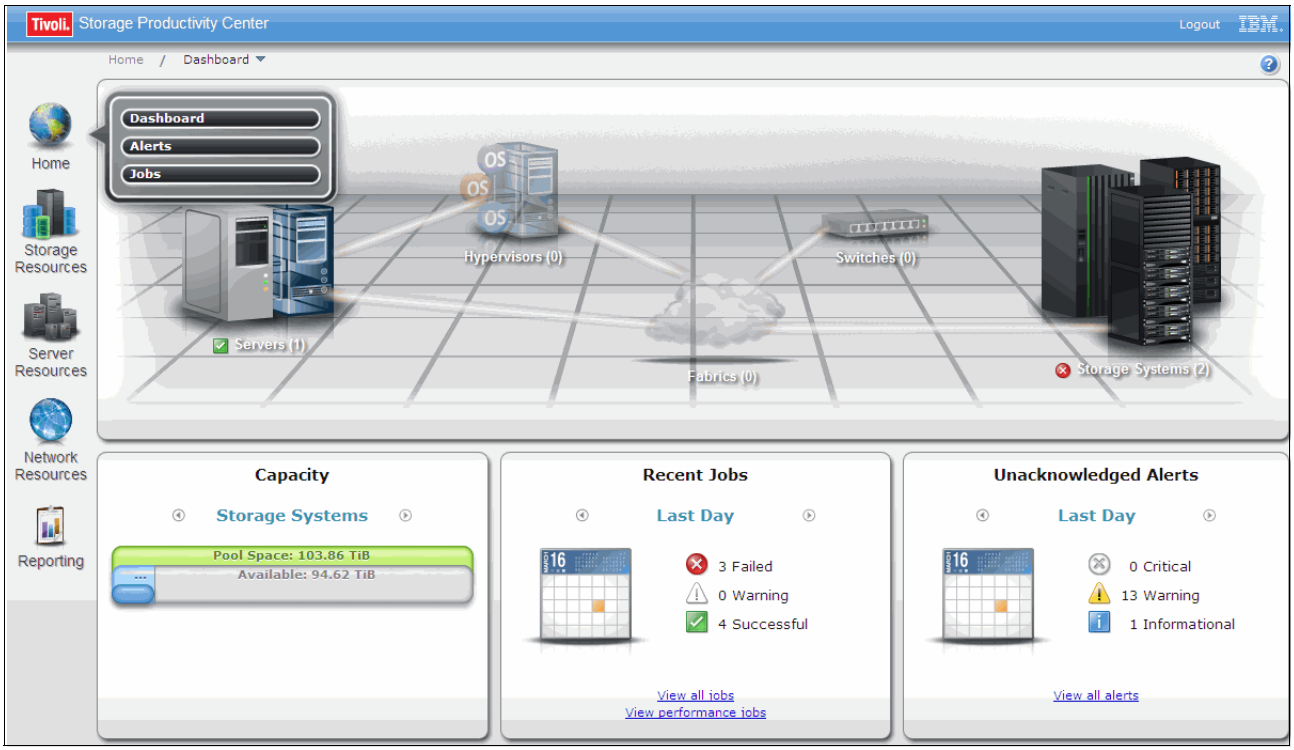


Figure 7-86 Tivoli Storage Productivity Center GUI Dashboard

The icons on the left side of the display provide links to groups of related resources. The XIV Storage System details are under the Storage Resources icon. See Figure 7-87. The overview window can be easily customized to show the information that is most relevant to your requirements.

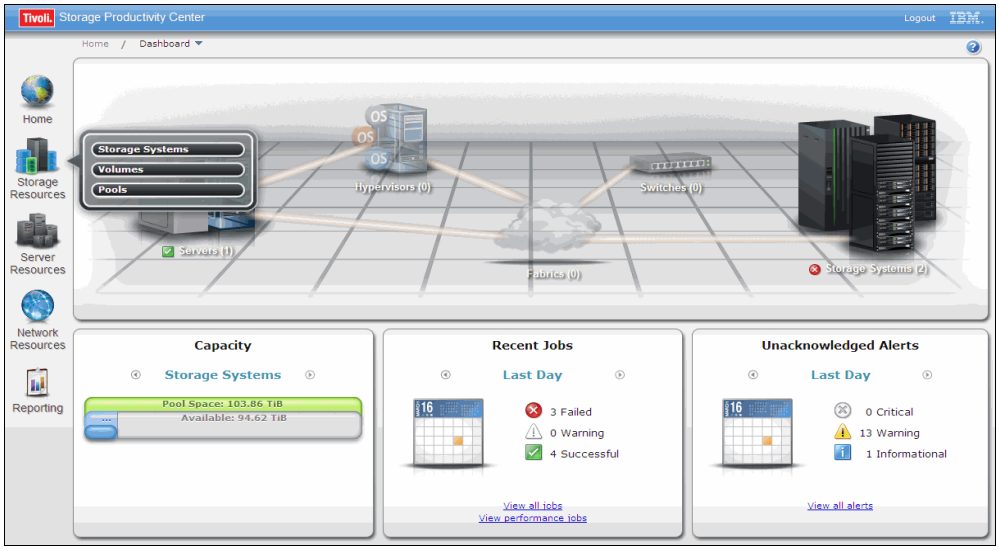


Figure 7-87 Tivoli Storage Productivity Center GUI storage resources

Selecting **Storage Systems** displays a list of the available storage systems. Double-click the name of a storage system and the Overview window is displayed. See Figure 7-88.

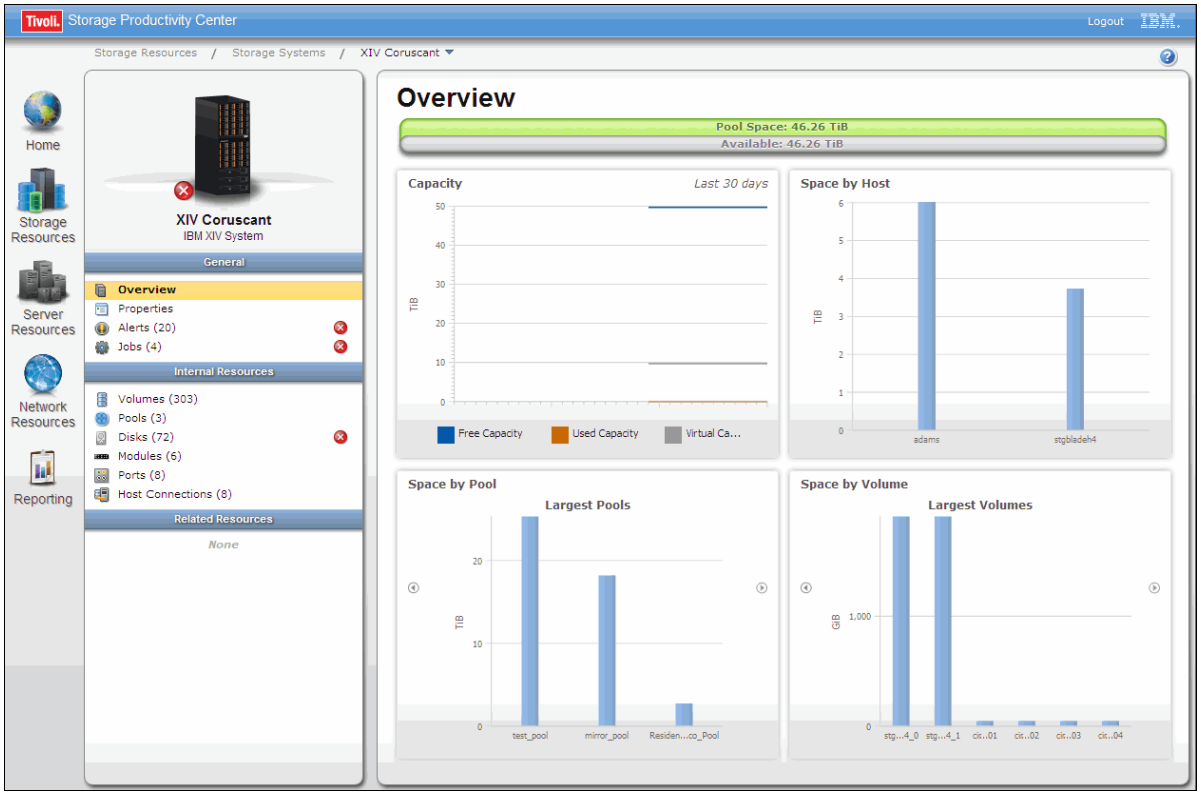


Figure 7-88 Tivoli Storage Productivity Center GUI resource overview

Figure 7-89 shows the options that are available from the drop-down menu in each quadrant of the display.

From the Overview window, you can also navigate to information about the XIV Storage System. The following links are provided:

- Volumes: Volume status and capacity
- Pools: Pool status and capacity
- Disks: Status of the XIV disks by module
- Modules: XIV module status
- Ports: Port status and worldwide port name (WWPN)
- Host connections: Defined host connections and volume mappings

The Overview window is also available from the Alerts display. If an alert is related to an XIV Storage System, the name of the resource associated with the event is a link to the Overview window. In this way, details about the resource are easily available for investigation into the cause of a specific alert.

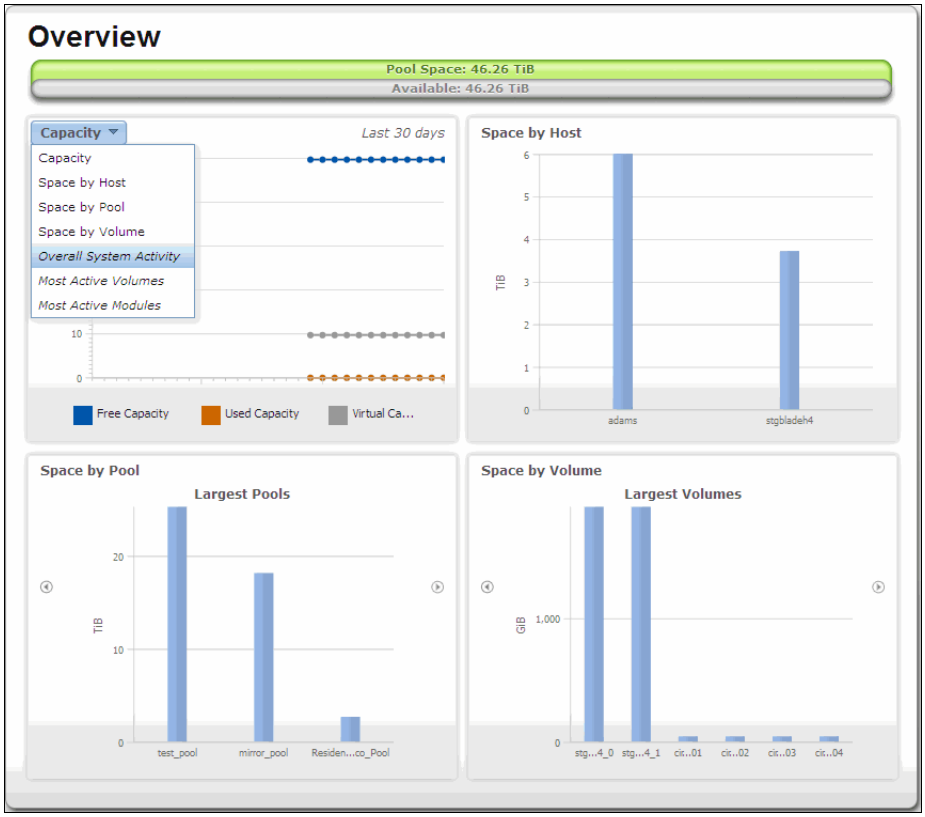


Figure 7-89 Tivoli Storage Productivity Center GUI overview customization

7.8 Custom monitoring

The previous sections of this chapter focused on the predefined methods of monitoring the XIV Storage System. We now look at ways to combine the various commands, options, and features so that you customize the monitoring of the XIV Storage System based on your specific needs or requirements.

Combining the power of XCLI with the flexibility of scripting languages creates endless possibilities and ways to monitor the system. For example, you can create a script to monitor a specific component and alert you when it exceeds a custom threshold that you define.

7.8.1 Custom monitoring of a volume

Example 7-47 demonstrates how to create a Windows Scripting Host (WSH) script to monitor the capacity of a specific volume and issue a custom event to alert you when the used capacity reaches a predefined limit. The script can be scheduled to run automatically at specific intervals using the standard Windows Task Scheduler.

Example 7-47 WSH script to monitor a volume

```
<job id="XIV_Vol_Monitor">
<script language="JScript">

//WSH Example for Custom Monitoring of an XIV Volume
//Modify the following parameters with proper values

   var IPAddr = "192.168.1.150";
   var Volume = "itso_vol_2";
   var Threshold = 700;
   var Severity = "Warning";
   var Username = "admin";
   var Password = "adminpass";

//Main Script Starts Here
   var CmdString = "";
   var String1 = "";
   var output = "";

   var xclicmd = "xcli -m " + IPAddr + " -u " + Username + " -p " + Password + "
";
   var Command = "vol_list vol=" + Volume + " -t used_capacity";

   CmdString = xclicmd + Command;

   var WshShell = WScript.CreateObject("WScript.Shell");
   var oExec = WshShell.Exec("%comspec% /d/q");

   oExec.StdIn.WriteLine("prompt $_");
   oExec.StdIn.WriteLine("cd \\");
   oExec.StdIn.WriteLine(CmdString);
   oExec.StdIn.WriteLine("exit");

   while (!oExec.StdOut.AtEndOfStream)
   {
```

```
      String1 = oExec.StdOut.ReadLine() + "\n";
      String1 = String1.replace(/^\s+|=|\s+$/g,"");
      if (String1 != "")
        output = String1;
   }

   if (output >= Threshold)
   {
      Command = "custom_event description='Vol " + Volume + " exceeded threshold
of " + Threshold + " GB' severity=" + Severity;

      CmdString = xclicmd + Command;

      WshShell = WScript.CreateObject("WScript.Shell");
      oExec = WshShell.Exec("%comspec% /d/q");

      oExec.StdIn.WriteLine("prompt $_");
      oExec.StdIn.WriteLine("cd \\");
      oExec.StdIn.WriteLine(CmdString);
      oExec.StdIn.WriteLine("exit");
   }

   while (oExec.Status == 0)
      WScript.Sleep(100);

   WshShell.LogEvent(0, "XIV Vol Monitor Script Completed");

</script>
</job>
```

The script in Example 7-47 on page 407 begins with defining variables that are used throughout the process. It then builds an XCLI command that is based on these variables and runs the command in a shell.

The output from XCLI commands can be customized using the **-t** option to limit the display to only the specific fields in which we are interested. In this case, we need only the used capacity value of the volume. Therefore, the **-t** option is used to limit the output to only the `used_capacity` field, as shown in Example 7-48. This option makes the output of the XCLI command easy to parse.

Example 7-48 Limiting output of XCLI commands using the -t option

```
var Command = "vol_list vol=" + Volume + " -t used_capacity"
```

You can get a complete list of the field options available for any command by using the XCLI **help** option with the **format=full** parameter, as shown in Example 7-49.

Example 7-49 Using the XCLI help command to list available output fields

```
>> help command=vol_list format=full
```

The output from the XCLI command is parsed and compared with the predefined custom threshold. If the threshold is exceeded, an XCLI command is used to create a custom event with a custom description and severity level, as shown in Example 7-50. An event rule can be used on the XIV Storage System to notify a user of the condition.

Example 7-50 Using XCLI to generate a custom event

```
if (output >= Threshold)
   {
      Command = "custom_event description='Vol " + Volume + " exceeded threshold
of " + Threshold + " GB' severity=" + Severity;
```

After the script is created, use the Windows Task Scheduler to schedule the script to run automatically on a daily basis by clicking **Start** → **Programs** → **Accessories** → **System Tools** → **Scheduled Tasks**, as shown in Figure 7-90.

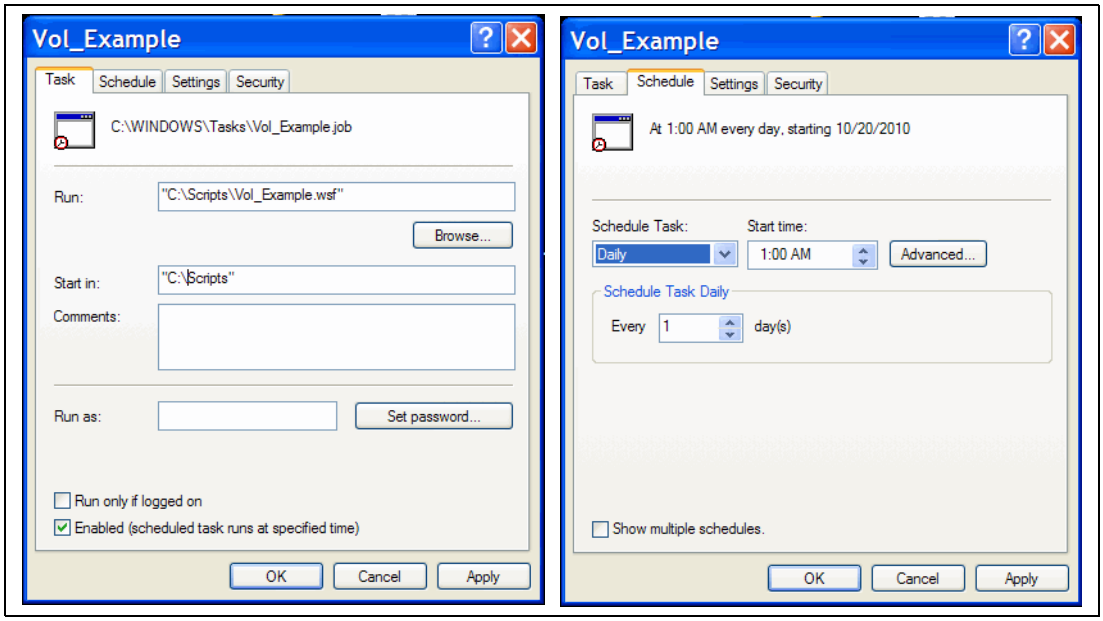


Figure 7-90 Using Windows Task Scheduler to run custom monitoring scripts

After running this script, a custom event is generated if the used capacity threshold is exceeded. You can see the custom event in the XIV Storage System event log, as shown in Figure 7-91. An event notification rule can be defined based on severity or the CUSTOM_EVENT event code to send an email or SNMP message, as described in 7.2, “Monitoring using the IBM XIV Mobile Dashboard” on page 351.

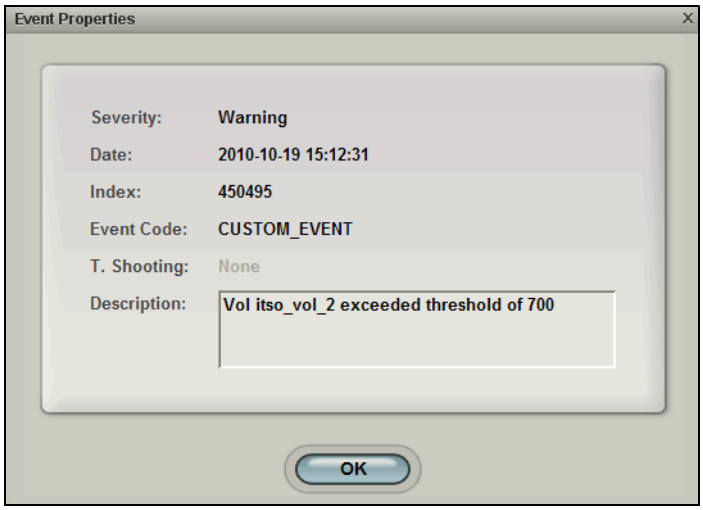


Figure 7-91 A custom event generated by the WSH script using XCLI

7.8.2 Custom monitoring of performance attributes

In Example 7-51, we demonstrate how to create a **WSH** script to monitor system performance, specifically the write hit latency, and issue a custom event to alert you when the latency reaches a predefined limit. The script can be scheduled to run automatically at specific intervals using the standard Windows Task Scheduler.

Example 7-51 WSH script to monitor write hit latency

```
<job id="XIV_Latency_Monitor">
<script language="JScript">
//WSH Example for Custom Monitoring of an XIV Latency Performance
//Modify the following parameters with proper values
   var IPAddr = "192.168.1.150";
   var Threshold = 3000;
   var Severity = "Warning";
   var Username = "admin";
   var Password = "adminpass";

//Main Script Starts Here
   var strValidChars = "0123456789";
   var strChar = "";
   var CmdString = "";
   var String1 = "";
   var output = "";
   var Exceeded = 0;

   var xclicmd = "xcli -m " + IPAddr + " -u " + Username + " -p " + Password + "
";

   var now = new Date();
   var EndDate = now.getYear() + "-" + (now.getMonth()+1) + "-" +
(now.getDate()-1) + ".00:00:00";

   var Command = "statistics_get count=1440 interval=1 resolution_unit=minute -t
write_hit_medium_latency end=" + EndDate;

   CmdString = xclicmd + Command;

   var WshShell = WScript.CreateObject("WScript.Shell");
   var oExec = WshShell.Exec("%comspec% /d/q");

   oExec.StdIn.WriteLine("prompt $_");
   oExec.StdIn.WriteLine("cd \\");
   oExec.StdIn.WriteLine(CmdString);
   oExec.StdIn.WriteLine("exit");

while (!oExec.StdOut.AtEndOfStream)
   {
      String1 = oExec.StdOut.ReadLine() + "\n";
      String1 = String1.replace(/^\s+|=|\s+$/g,"");
      if (String1.length == 0) continue;

      strChar = String1.charAt(0);
      if (strValidChars.indexOf(strChar) == -1) continue;
      if (String1 >= Threshold) Exceeded = 1;
```

```
   }
      if (Exceeded)
   {
      Command = "custom_event description='Medium Latency exceeded threshold of "
+ Threshold + " during 24 hour period ending " + EndDate + "' severity=" +
Severity;
      CmdString = xclicmd + Command;
      WshShell = WScript.CreateObject("WScript.Shell");
      oExec = WshShell.Exec("%comspec% /d/q");
      oExec.StdIn.WriteLine("prompt $_");
      oExec.StdIn.WriteLine("cd \\");
      oExec.StdIn.WriteLine(CmdString);
      oExec.StdIn.WriteLine("exit");
   }

   while (oExec.Status == 0)
      WScript.Sleep(100);

   WshShell.LogEvent(0, "XIV Latency Monitor Script Completed Successfully");

</script>
</job>
```

The script in Example 7-51 on page 410 also begins with defining variables that are used throughout the process. It then builds an XCLI command that is based on these variables and runs the command in a shell.

The output from XCLI is again customized by using the **-t** option to limit the display to only the specific fields in which we are interested. In this case, we need only the write hit medium latency field. This option makes the output of the XCLI command easy to parse.

The **XCLI** command is structured in a way to provide output during a custom window of time. In Example 7-52, we gather data from every minute during a 24-hour period on the previous day that the script is run.

Example 7-52 Limiting the output of XCLI commands using the -t option

```
var Command = "statistics_get count=1440 interval=1 resolution_unit=minute -t
write_hit_medium_latency end=" + EndDate;
```

In Example 7-53 on page 411, each resulting value is compared to the threshold. If any one of them exceeds the value, it sets a flag to be used to generate a custom event.

Example 7-53 Looping through the XCLI command output

```
while (!oExec.StdOut.AtEndOfStream)
   {
      String1 = oExec.StdOut.ReadLine() + "\n";
      String1 = String1.replace(/^\s+|=|\s+$/g,"");
      if (String1.length == 0) continue;

      strChar = String1.charAt(0);
      if (strValidChars.indexOf(strChar) == -1) continue;
      if (String1 >= Threshold) Exceeded = 1;
   }
```

In Example 7-54, we see again how to generate a custom event when our wanted threshold is exceeded.

Example 7-54 Using XCLI to generate a custom event

```
if (Exceeded)
   {
      Command = "custom_event description='Medium Latency exceeded threshold of "
+ Threshold + " during 24 hour period ending " + EndDate + "' severity=" +
Severity;
```

After running this script, you can see the custom event in the XIV Storage System event log, as shown in Figure 7-92.

After running this script, a custom event is generated if the write hit medium latency threshold is exceeded. You can see the custom event in the XIV Storage System event log, as shown in Figure 7-92. An event notification rule can be defined based on severity or the CUSTOM_EVENT event code to send an email or SNMP message, as described in 7.2, “Monitoring using the IBM XIV Mobile Dashboard” on page 351.

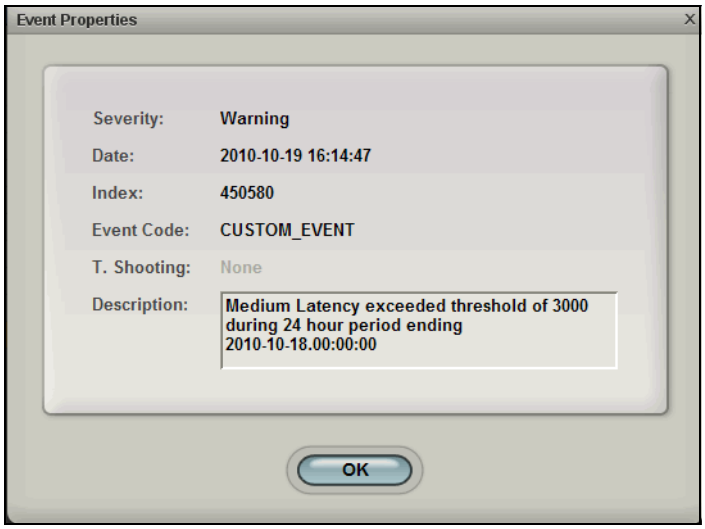


Figure 7-92 A custom event generated by WSH script by using XCLI

7.8.3 Summary

These simple examples are just a small subset of what can be accomplished by combining the power of XCLI commands with the flexibility of scripting.

By using the output from multiple commands or multiple fields, it is possible to monitor any combination of components or parameters of the system and act upon them based on criteria that you define.

Similar results can be accomplished on UNIX-based platforms by using shell scripting and **crontab** for scheduling custom monitoring processes.

7.9 Microsoft System Center Operations Manager

Microsoft System Center Operations Manager (SCOM) is the Microsoft end-to-end service-management product. From a single interface, you can verify the state, health, and performance of your Windows servers, certain UNIX-based computer systems, and other data center equipment, based on the availability of relevant management packs. SCOM can generate alerts when it identifies issues with availability, performance, configuration, or security.

Depending on your needs, using SCOM delivers the following benefits:

- Storage administrators can monitor multiple storage products from different vendors.
- Microsoft customers can monitor IBM systems using their familiar tools.
- System administrators can manage both storage systems and other IT systems from a single consistent interface.

A key feature of SCOM is its use of *management packs,* which contain application or product-specific rules and filters. Vendors can create their own management packs relevant to their own products. The IBM Storage Management Pack for SCOM is actually a set of management packs that allow you to access and monitor the following IBM storage systems using Microsoft SCOM:

- IBM System Storage DS8000
- IBM Storwize V3500 (since Version 2.1)
- IBM Storwize V3700 (since Version 2.1)
- BM Storwize V5000 (since Version 2.2)
- IBM Storwize V7000
- IBM Flex System V7000 (since Version 1.3)
- IBM System Storage SAN Volume Controller
- IBM XIV Storage System

The rules provided by the management packs remove a significant amount of work when implementing a monitoring solution. SCOM offers methods to monitor four areas of focus:

- Availability
- Configuration
- Performance
- Security

Currently, the IBM Storage Management Pack focuses on availability.

7.9.1 Prerequisites

To monitor your XIV Storage System with SCOM, you must meet the following prerequisites:

- Microsoft System Center Operations Manager 2007 R2, Microsoft System Center Operations Manager 2012, Microsoft System Center Operations Manager 2012 SP1 or Microsoft System Center Operations Manager 2012 R2
- IBM Storage Management Pack v1.1, v1.2, v1.3, V2.1, or V2.2 for SCOM
- XIV Storage System Software v10.2.2 or later

The XIV Storage Management GUI or XCLI does *not* need to be installed on the SCOM server. The IBM Storage Management Pack installs all the software that is required for SCOM to access the XIV Storage System.

7.9.2 Installing SCOM and the IBM Storage Management Pack

Installing SCOM onto a Windows server is not a trivial task and is beyond the scope of this publication. Assuming that SCOM is already installed and operating correctly, this section shows you how to install the IBM Storage Management Pack for SCOM. The release notes and installation instructions are available on the “Select fixes: Enterprise Storage Servers, XIV Storage System” web page:

http://ibm.co/1zpWKSH

You can also see the Microsoft System Center Operations Manager website:

http://www.microsoft.com/en-us/server-cloud/system-center/operations-manager.aspx

To install the IBM Storage Management Pack for SCOM, complete the following steps:

1. Download the relevant IBM Management Pack to the SCOM server (x86 or x64). After it is downloaded, double-click **IBM_Storage_MP_for_SCOM-windows** to start the installation.
2. If the IBM Storage Solutions External Runtime Components are not installed, you are prompted to install them. This package installs a version of Python called $xPYV$, which is used by the IBM Storage Management Pack for scripting. It does not interfere with any versions of Python you already have installed.
3. After the Runtime Components have been installed, the Storage Management Pack InstallationShield wizard now opens. Click **Next** and you are prompted to accept the IBM license.
4. You are now prompted to either do a complete installation or a custom installation. Because the IBM Storage Management Pack consists of nine management packs, you can choose which packs to install. The list of packs is shown in Figure 7-93 on page 414. The size of each pack is trivial (less than 200 KB each).

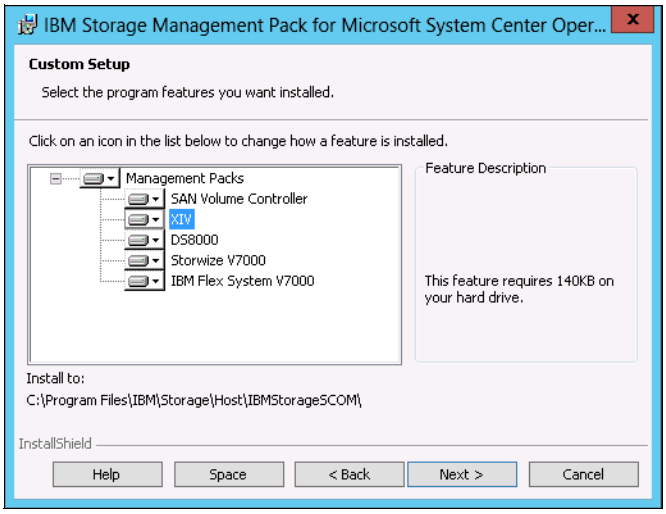


Figure 7-93 IBM Management Pack custom installation

5. Follow the prompts to complete the installation.

   You will later notice that the installation of the IBM Storage Management Pack also includes a stand-alone command-line interface (CLI) utility, `scomu.cmd`, which you can use from a desktop shortcut or from the Run command box in the Windows Start menu. We will be using a set of CLI commands to configure the IBM storage systems that are monitored by Microsoft SCOM, as explained in the next section.

7.9.3 Importing the management packs and Adding IBM XIV systems

The installation of the IBM Storage Management Pack does not import the management packs themselves into SCOM. You must import each specific management pack (for each storage system type). Complete the following steps:

1. In the SCOM Administration window, right-click **Management Packs** and select **Import Management Packs** (see Figure 7-94) to open the Import Management Packs window.

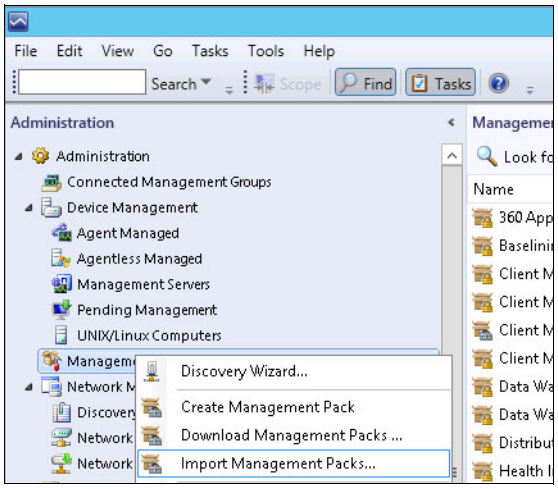


Figure 7-94 Importing management packs into SCOM

2. Click **Add**, and then click **Add from disk**. An online catalog connection message is displayed. Click **No** to locate the management pack locally.

   When the Select Management Packs to Import window opens, select the `C:\ProgramFiles\IBM\Storage\Host\IBMStorageSCOM\mps` directory, and then select the following two files:

   - `IBM.Storage.Common.mp`
   - `IBM.Storage.XIV.mp`

3. When the files are selected, click **Open**. The Import Management Packs window now lists the Management Packs to be added.
4. Click **Install** to start the import, as shown in Figure 7-95. When the Management Packs are successfully imported, click **Close**.

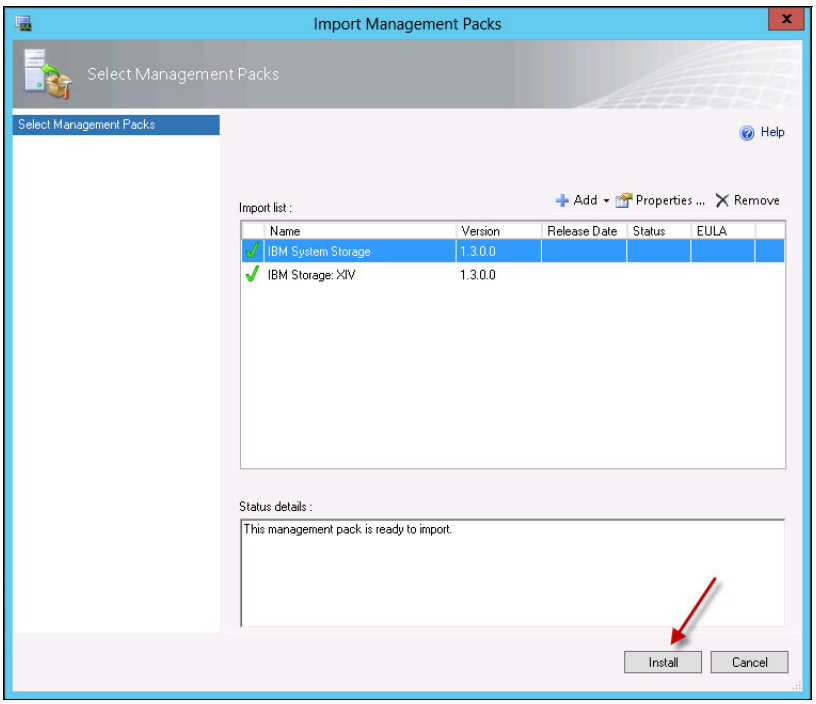


Figure 7-95 Install the Management Packs

You have now imported the Management Packs into SCOM.

5. The installation of the Storage Management Pack creates an icon on the desktop entitled *IBM Storage SCOM-control Utility Command Prompt*.

   Doubling-clicking this icon opens a Windows command prompt in the correct directory to use the IBM provided **scomu** command. You now need to use the **scomu** command from within this command prompt to add the management server. Use the following syntax for this command:

   ```
   scomu --sc-set --servername localhost
   ```

   Tip: Pay close attention to the syntax. You must use the double dash where required or the command fails. So, for example, **--servername** works, but **-servername** does not.

   An example of the syntax you use and the expected responses is shown in Example 7-55 on page 416. In this example, the management server is the machine running the SCOM platform and not another server running the SCOM agent.

   Example 7-55 Adding the management server

   ```
   scomu.cmd --sc-set --servername localhost
   Checking the connection to the management server...
   The connection to the management server is OK
   The connection to the management server is set
   ```

6. From the IBM Storage SCOM-control Utility Command prompt you now need to add your IBM XIV systems to the monitoring list. Use the following syntax for this command:

   scomu --add -t xiv --ip <*ip address*> --username <*username*> --password <*password*>

Tip: Pay close attention to the syntax. You must use the double dash where required or the command fails. So, for example, **--password** works, but **-password** does not.

An example of the syntax you use and the expected responses is shown in Example 7-56. The device ID is the serial number of the XIV Storage System. In Example 7-56, the device ID is 10114 because the XIV Storage System serial number is 1310114. The user name and password combination must be for an existing user name defined to the XIV Storage System. If a read-only user is used, mirroring is not monitored.

Example 7-56 Adding an XIV Storage System to the SCOM monitoring list

```
scomu --add -t xiv --ip 10.0.20.102 --username itso --password password
Connecting to the device ...
1 IBM XIV Storage System is found.
device ID: 10114, code level: 11.5.0-esp3-p20140623_181834

The connection is OK.
Trying to add the connection...
The connection has been added.
```

7. The **scomu --add** command allows you to only define one XIV Storage System IP address at a time. Repeat the **--add** command for each XIV Storage System management IP address. You must run the **add** command three times for each XIV Storage System.

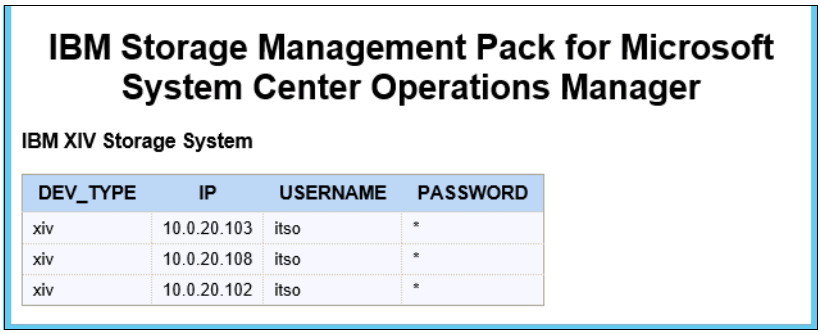
Tip: If you have clustered SCOM servers, you must configure each XIV Storage System and each XIV Storage System management IP address on each server, as well as the Root Management Server in the Management Group. This action ensures that monitoring continues to work after a failover.

8. Because you defined your IBM XIV systems, you can list them and confirm that they are all defined by running **scomu --list**, as shown in Example 7-57.

Example 7-57 Listing the monitoring systems

```
scomu --list
3 connections have been found.
```

The default output format is HTML, and the list is displayed in the default web browser, as shown in Figure 7-96 on page 417.


IBM Storage Management Pack for Microsoft System Center Operations Manager

IBM XIV Storage System

DEV_TYPE	IP	USERNAME	PASSWORD
xiv	10.0.20.103	itso	*
xiv	10.0.20.108	itso	*
xiv	10.0.20.102	itso	*

Figure 7-96 Output of the scomu --list command

Adding IBM XIV systems to the monitoring list is complete. You now need to configure the SCOM Management Pack.

7.9.4 Configuring the SCOM Management pack

SCOM monitors various different constructs, such as alerts, events, systems, and components. Each construct has a data collection interval that ranges from 300 seconds to 1800 seconds (the time varies based on the construct). This situation means that if you add an XIV Storage System to SCOM using the **`scomu --add`** command, there might be a delay before the XIV Storage System is displayed in any of the views.

Changing the default discovery intervals

To change the default discovery intervals, complete the following steps:

1. In the left pane of the Operations console, click **Authoring** → **Management Pack Objects** → **Object Discoveries**.
2. Click **XIV Configuration Discovery**, right-click, and click **Overrides** → **Override the Object Discovery** → **For all objects of class: Windows Computer**.

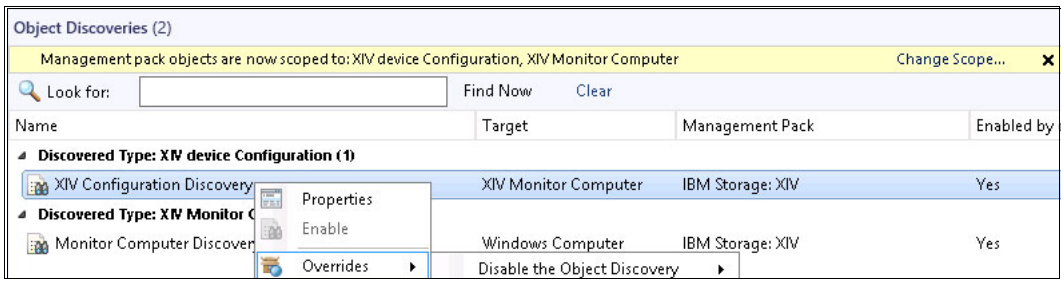


Figure 7-97 SCOM Object Discovery settings

3. From the Override Properties window, select the **Override** check box in the Frequency in seconds row. You can now change the **Override Value** in seconds from 14400 to a different value. After having selected the Default Management Pack as destination Management Pack, you can click **OK** to apply the change.

Selecting a shorter monitoring interval

At first glance, a monitoring interval of 10 minutes appears to be long. If you decide to reduce the intervals, you must decide which interval values to use. Although a long monitoring interval might sound good, additional monitoring increases the workload of both the monitored devices and of the SCOM server. This situation might lead to unforeseen consequences.

One method to determine how quickly SCOM is able to complete a scan of your devices. The default log folder for SCOM is `C:\Program Files\IBM\Storage\Host\IBMStorageSCOM\log`. In that folder, you see two files for each device type. The XIV Storage System uses the following log files:

scom_xiv.log This log file is updated when SCOM checks the state of each monitored XIV Storage System component.

scom_xiv_event.log This log file is updated when SCOM collects new events from the XIV Storage System event logs.

In Example 7-58, the event log for an XIV Storage System is examined in about 3 seconds. In this example, one new event is identified.

Example 7-58 The scom_xiv_event log

```
2013-09-24 13:37:38 INFO     Enter collect()
2013-09-24 13:37:39 INFO     Connecting to 10.0.20.102
```

```
2013-09-24 13:37:40 INFO     1 IBM XIV Storage System is found.
device ID: 10114, code level: 11.4.0-esp4-p20130815
2013-09-24 13:37:41 INFO     Enter get_xiv_events()
2013-09-24 13:37:41 INFO     Enter get_begin_time_idx()
2013-09-24 13:37:41 INFO     Exit get_begin_time_idx()
2013-09-24 13:37:41 INFO     begin_time = 2013-09-24.13:26:37, begin_idx = 8971
2013-09-24 13:37:41 INFO     event_list after 2013-09-24.13:26:37
2013-09-24 13:37:41 INFO     adding 1 new events to SCOM
2013-09-24 13:37:41 INFO     Enter write_end_time_idx()
2013-09-24 13:37:41 INFO     Exit write_end_time_idx()
2013-09-24 13:37:41 INFO     end time = 2013-09-24.13:29:18, last index = 8972
2013-09-24 13:37:41 INFO     Exit get_xiv_events()
2013-09-24 13:37:41 INFO     Finished collecting events from xiv 10114
successfully
2013-09-24 13:37:41 INFO     Return items to scom
2013-09-24 13:37:41 INFO     Exit collect()
```

In Example 7-59, the discovery process for monitoring is run against the same XIV Storage System in about 16 seconds.

Example 7-59 The scom_xiv log

```
2013-09-24 13:30:09 INFO     Enter discover_device_by_configuration()
2013-09-24 13:30:09 INFO     Connecting to 10.0.20.102
2013-09-24 13:30:11 INFO     1 IBM XIV Storage System is found.
device ID: 10114, code level: 11.4.0-esp4-p20130815
2013-09-24 13:30:11 INFO     Started discovering 10114 ...
2013-09-24 13:30:11 INFO     Enter get_xiv_devices()
2013-09-24 13:30:12 INFO     Exit get_xiv_devices()
........
2013-09-24 13:30:14 INFO     Enter get_xiv_host_mapping()
2013-09-24 13:30:16 INFO     XIV host url 10.0.20.102 discovered
2013-09-24 13:30:16 INFO     XIV(10.0.20.102) can be connected.
2013-09-24 13:30:17 INFO     Exit discover_configuration()
2013-09-24 13:30:18 INFO     Exit get_xiv_host_mapping()
2013-09-24 13:30:18 INFO     Finished discovery for xiv 10114 successfully
2013-09-24 13:30:18 INFO     Returning data to SCOM
2013-09-24 13:30:25 INFO     Exit discover_device_by_configuration()
```

As you add additional devices, the total time to complete a scan increases because the scans are performed serially. Allow a significant buffer of time between the completion of one scan and the start of the next scan.

7.9.5 Monitoring your XIV Storage System with SCOM

The IBM Storage Management Pack monitors each XIV Storage System for what it classifies as *events* and *alerts*. These items are updated every 600 seconds. In addition, various components of each XIV Storage System are monitored, and the status is updated every 300 seconds. Components can be in one of three states:

- Healthy
- Warning
- Error

Alerts

Alerts are events that need to be investigated. They indicate that an application or device needs attention. They can be thought of as a to-do list for an administrator. Alerts are generated in two ways:

- By monitoring the event logs of each managed device.
- By monitoring the health state of each component. For example, if the XIV Storage System marks a module or disk as *Failed*, SCOM raises a Critical alert. If a component changes from Failed to Ready, SCOM raises a Warning alert. Examples of alert monitors include Disk Monitor, Module Monitor, and XIV System Monitor.

Tip: If a component fails and is then repaired (for example, when a disk fails and is then replaced with a new disk), the resolution state of the Critical alert is automatically changed by SCOM from *New* to *Closed*. However, new alerts might be created with a severity of Warning, which indicates that the status is now Ready.

Each alert appears in IBM XIV Systems Alerts tab with the Resolution State *New*. Change it manually to *Closed,* if necessary, as shown in Figure 7-98.

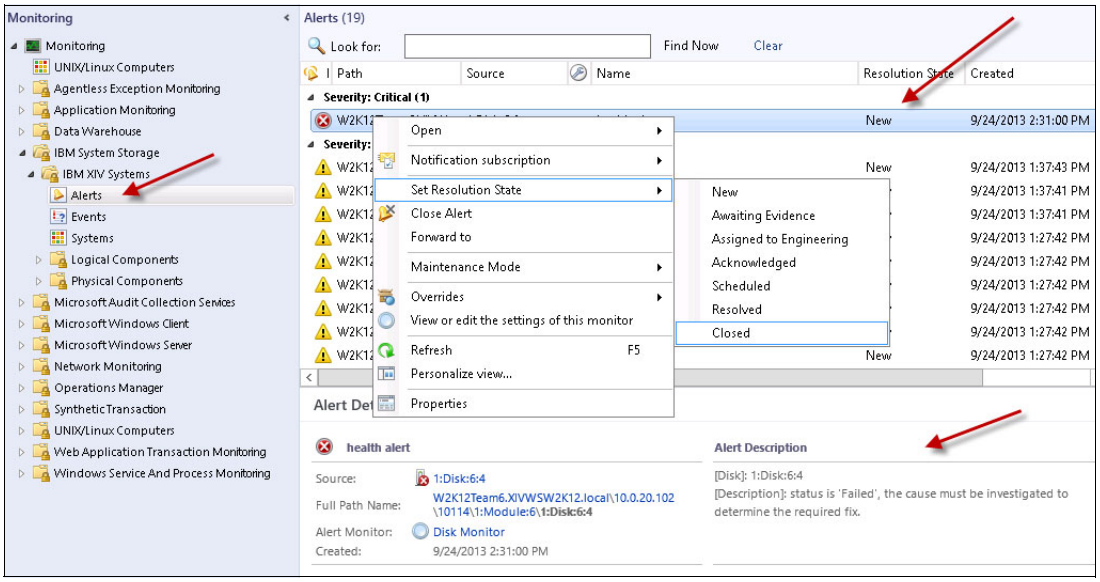


Figure 7-98 SCOM Alerts window

Events

The Events window shows events for each monitored XIV Storage System. SCOM places XIV Storage System events into three categories:

- Information
- Error
- Warning

The list of events is refreshed every 10 minutes. Events do not have nor need a resolution status. The initial event collection that SCOM runs does not collect events that are more than two days old. There is no way to change this setting. After the initial collection of events, SCOM collects only events created since the last event collection, up to a maximum of 300 events per collection. SCOM then examines the new events to see if alerts need to be raised. A short delay might occur between the collection of the events and the creation of the alert.

Systems

SCOM lists all defined IBM XIV systems that are being monitored in the Systems window. By default, the systems monitoring list is refreshed every 30 minutes. From the Systems view, it is possible to confirm the health of all the monitored components of each XIV Storage System and collect useful information.

In Figure 7-99, two IBM XIV systems are being monitored.

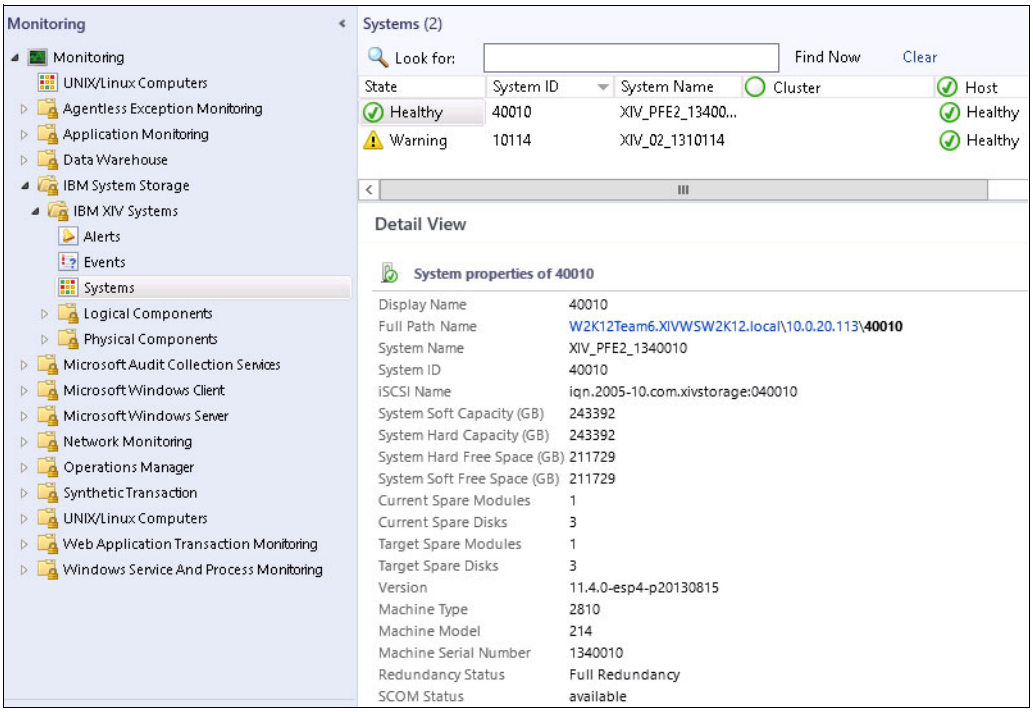


Figure 7-99 SCOM Systems view

Monitors

The following XIV Storage System components are monitored by SCOM. Each monitor is refreshed every 300 seconds.

- Logical components: These components are user-defined constructs:
  - Host Mappings
  - Mirrorings (This component is not monitored if the XIV Storage System user that was used to define the XIV Storage System to SCOM has only read-only privileges.)
  - Storage Pools
  - Volumes
- Physical components: These components represent XIV Storage System hardware or XIV Storage System definitions that relate to hardware. Hosts and clusters appear in this list because they represent physical hardware:
  - Clusters
  - Disks
  - Fibre Channel Ports
  - Hosts
  - IP Interfaces
  - iSCSI Ports
  - Modules

Diagram view

It is possible to show any of the views from SCOM as a diagram by right-clicking an object and clicking **Open** → **Diagram View**. For example, a failed disk can be displayed as a subcomponent of a module, as shown in Figure 7-100, where disk 12 on module 7 has failed. All other components of module 7 are currently healthy and are gathered together.

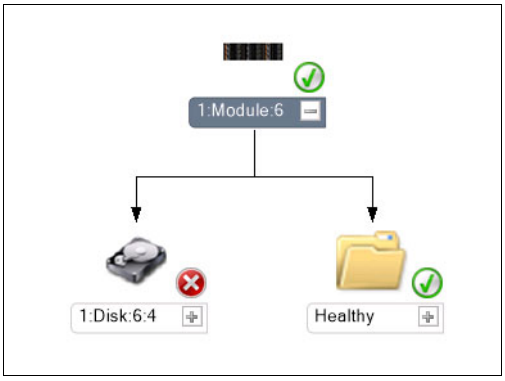


Figure 7-100 SCOM diagram view

7.9.6 Upgrading the IBM Storage Management Pack

If a new version of the IBM Storage Management Pack for Microsoft System Center Operations Manager is released, check the release notes to confirm any changes that you need to perform. An upgrade from Version 1.3 to 2.1 requires you to perform these tasks:

1. For the upgrade from Version 1.3 to Version 2.1, you need to delete your IBM XIV systems by using the **scomu** command. Use the following syntax for the command:

   ```
   scomu --del -t xiv --ip <ip adress>
   ```

2. Before the installation of Version 2.1, you need to delete the old XIV management pack in SCOM first, as shown in Figure 7-101.

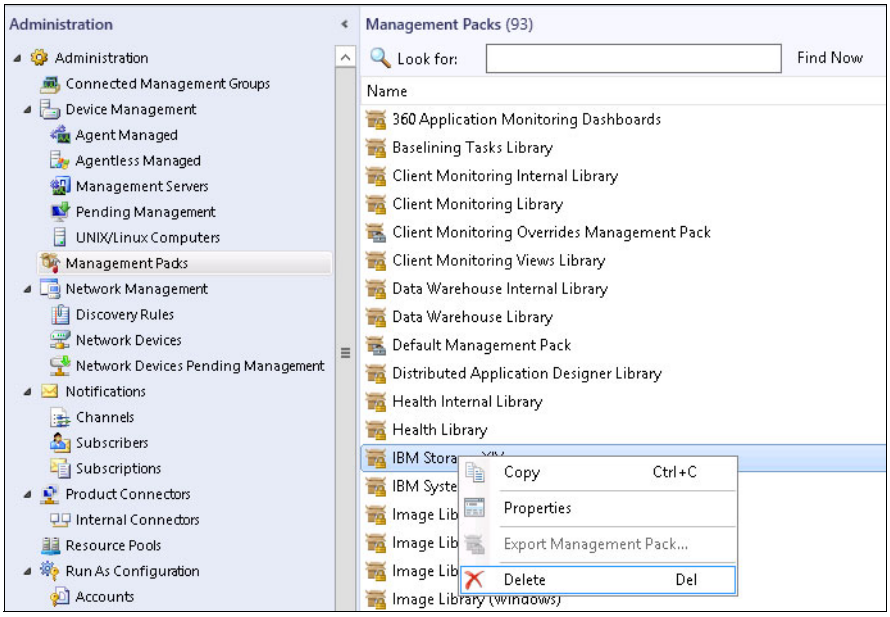


Figure 7-101 Delete the XIV management pack

3. Reinstall the IBM Storage Management Pack by using the same process that you used to install it. See 7.9.2, “Installing SCOM and the IBM Storage Management Pack” on page 414. The upgrade wizard is automatically started.

4. Import the updated management packs as described in 7.9.3, “Importing the management packs and Adding IBM XIV systems” on page 415. In Figure 7-102, you can see that Version 1.3.0.0 is imported for IBM System Storage and that both management packs Version 2.1.0.0 are selected for import.

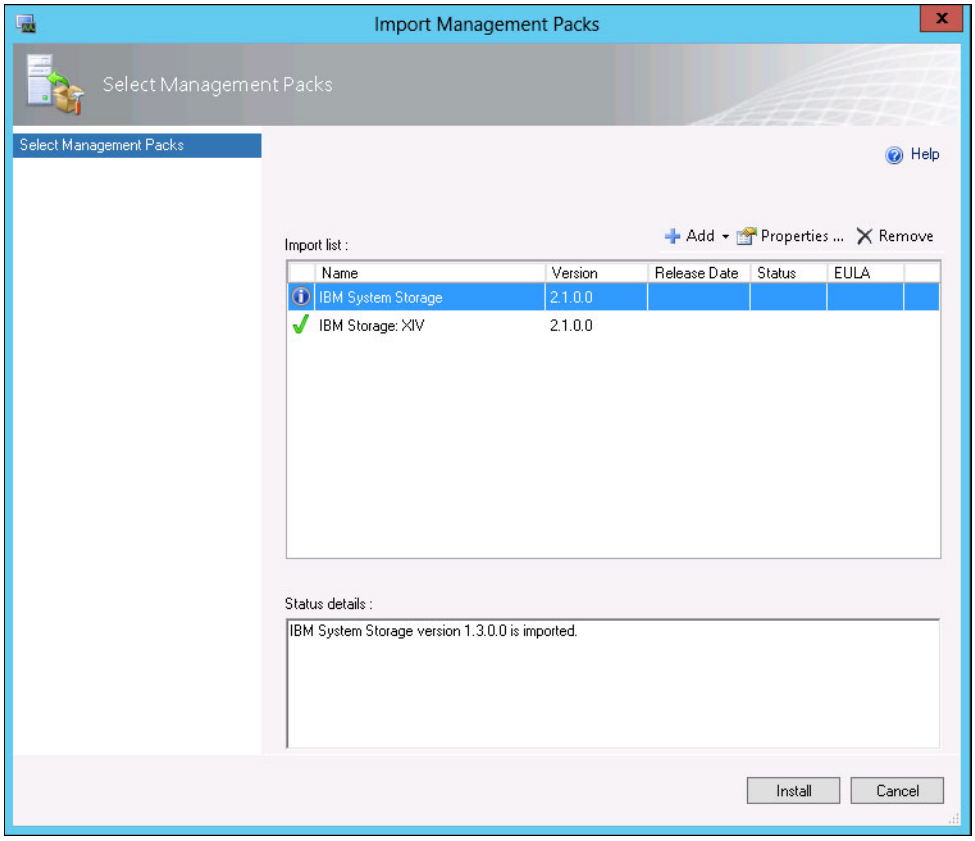


Figure 7-102 Upgrading the management packs

5. Since Version 2.1, you need to define the SCOM management server with the **scomu** command. This the syntax of the if you run it on the management server:

```
scomu.cmd --sc-set --servername localhost
```

6. Now, add the XIV systems again, as shown in Example 7-56 on page 417, with the following **scomu** command:

```
scomu --add -t xiv --ip <ip address> --username <username> --password
<password>.
```

A

Thin provisioning conceptual examples

In this appendix, we further explain thin provisioning principles by examining some basic examples. These examples incorporate all of the concepts inherent to the IBM XIV Storage System implementation of thin provisioning.

This appendix covers the following topics:

- System-level thin provisioning conceptual example
- Regular storage pool conceptual example
- Thinly provisioned storage pool conceptual example

© Copyright IBM Corp. 2014. All rights reserved.

System-level thin provisioning conceptual example

Figure A-1 shows the incremental allocation of capacity to both a regular storage pool and a thinly provisioned storage pool within the context of the global system soft and hard sizes. This example assumes that the soft system size has been defined to exceed its hard size. The deallocated capacity shown within the system's soft and hard space is represented by a discontinuity to convey the full scope of both the logical and physical view of the system's capacity. Each increment in the diagram represents 17 GB of soft or hard capacity.

When a regular storage pool is defined, only one capacity is specified, and this amount is allocated to the storage pool from *both* the hard and soft global capacity within the system.

When a thinly provisioned storage pool is defined, both the soft and hard capacity limits for the storage pool must be specified. These amounts are deducted from the system's global available soft and hard capacity.

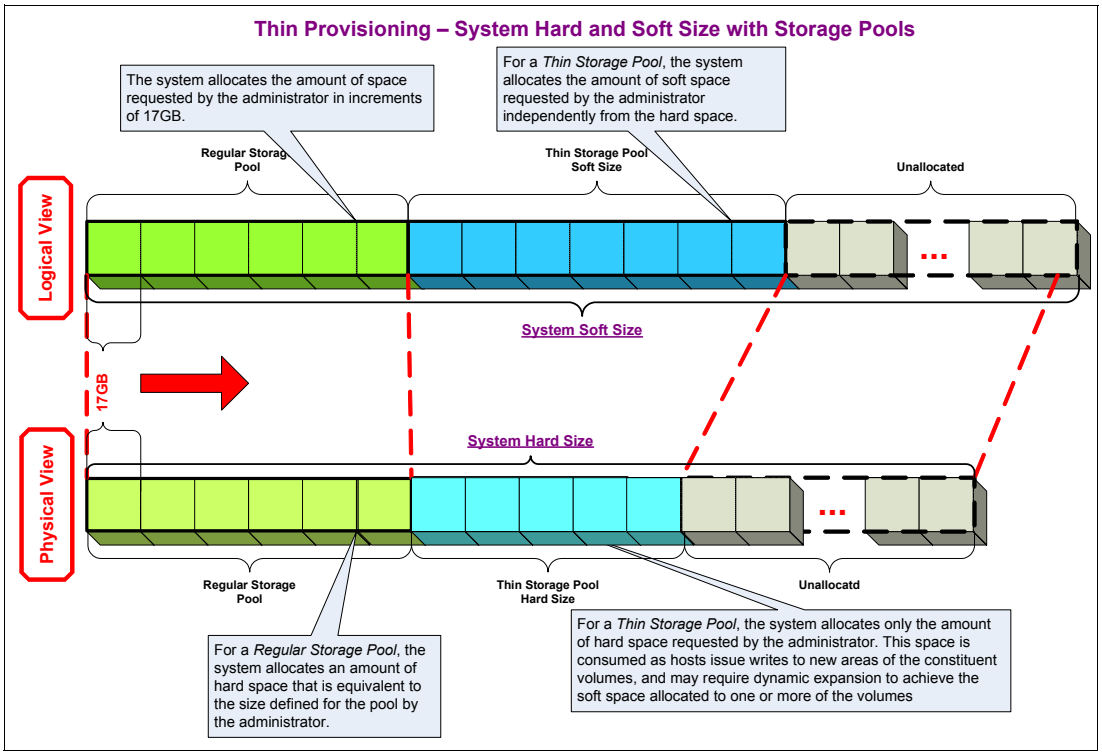

Thin Provisioning – System Hard and Soft Size with Storage Pools
The system allocates the amount of space requested by the administrator in increments of 17GB.
For a Thin Storage Pool, the system allocates the amount of soft space requested by the administrator independently from the hard space.
Regular Storag Pool
Thin Storage Pool Soft Size
Unallocated
Logical View
System Soft Size
17GB
System Hard Size
Physical View
Regular Storage Pool
Thin Storage Pool Hard Size
Unallocatd
For a Regular Storage Pool, the system allocates an amount of hard space that is equivalent to the size defined for the pool by the administrator.
For a Thin Storage Pool, the system allocates only the amount of hard space requested by the administrator. This space is consumed as hosts issue writes to new areas of the constituent volumes, and may require dynamic expansion to achieve the soft space allocated to one or more of the volumes

Figure A-1 Thin provisioning at the system level

Regular storage pool conceptual example

Next, Figure A-2 represents a focused view of the regular storage pool that is shown in Figure A-1 on page 426 and shows the division of both soft and hard capacity among volumes within the pool. The regular pool is the same size (102 GB) in both diagrams.

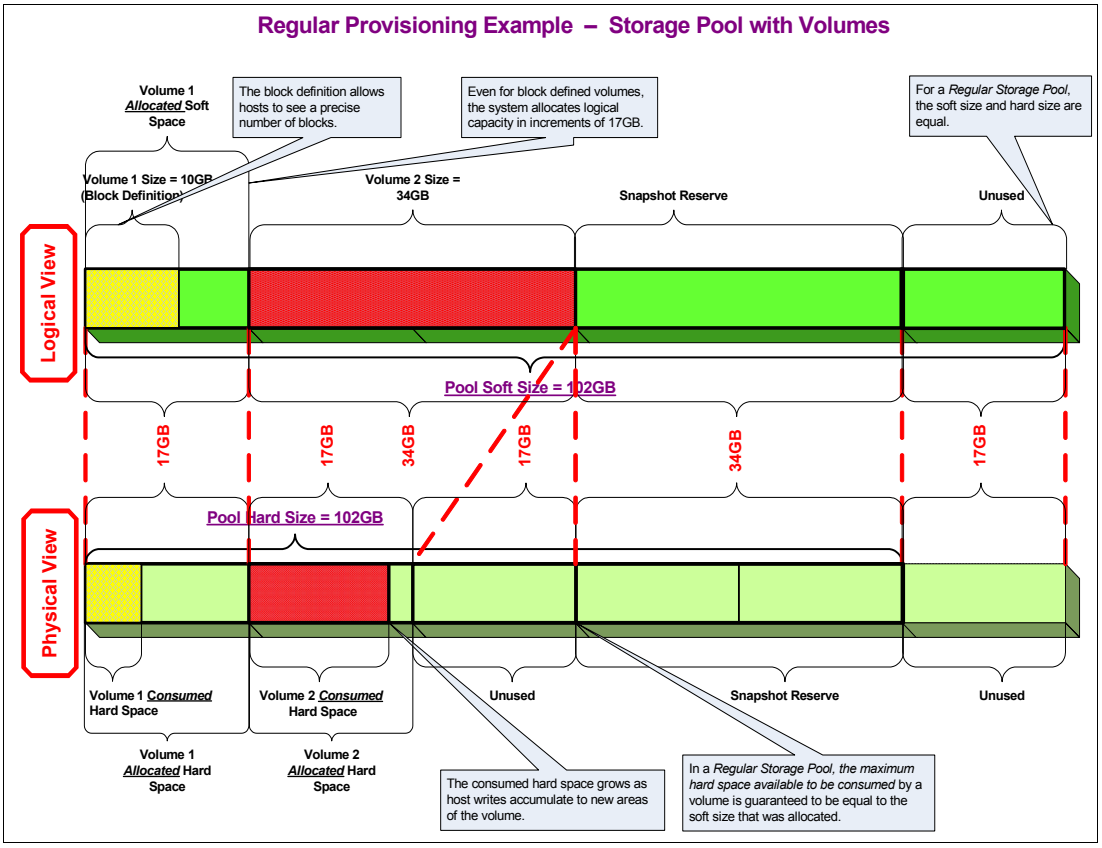

Regular Provisioning Example – Storage Pool with Volumes
Volume 1 Allocated Soft Space
The block definition allows hosts to see a precise number of blocks.
Even for block defined volumes, the system allocates logical capacity in increments of 17GB.
For a Regular Storage Pool, the soft size and hard size are equal.
Volume 1 Size = 10GB (Block Definition)
Volume 2 Size = 34GB
Snapshot Reserve
Unused
Logical View
Pool Soft Size = 102GB
17GB
17GB
34GB
17GB
34GB
17GB
Pool Hard Size = 102GB
Physical View
Volume 1 Consumed Hard Space
Volume 2 Consumed Hard Space
Unused
Snapshot Reserve
Unused
Volume 1 Allocated Hard Space
Volume 2 Allocated Hard Space
The consumed hard space grows as host writes accumulate to new areas of the volume.
In a Regular Storage Pool, the maximum hard space available to be consumed by a volume is guaranteed to be equal to the soft size that was allocated.

Figure A-2 Volumes and snapshot reserve space within a regular storage pool

Consider Volume 1. Although Volume 1 is defined as 19,737,900 blocks (10 GB), the soft capacity allocated is nevertheless composed of the minimum number of 17 GB increments needed to meet or exceed the requested size in blocks, which is in this case only a single 17 GB increment of capacity. The host, however, sees exactly 19,737,900 blocks. When Volume 1 is created, the system does *not* initially allocate any hard capacity. At the moment that a host writes to Volume 1, even if it is just to initialize the volume, the system allocates 17 GB of hard capacity. The hard capacity allocation of 17 GB for Volume 1 is shown in Figure A-2, although clearly this allocation is never fully used if the host-defined capacity remains only 10 GB.

Unlike Volume 1, Volume 2 has been defined in terms of gigabytes and has a soft capacity allocation of 34 GB, which is the amount that is reported to any hosts that are mapped to the volume. In addition, the hard capacity consumed by host writes has not yet exceeded the 17 GB threshold. Therefore, the system has so far *only* allocated one increment of 17 GB hard capacity. The hard capacity and the soft capacity allocated to a regular storage pool are equal by definition. Therefore, the remaining 17 GB of soft capacity assigned to Volume 2 is effectively preserved and remains available within the pool's hard space until it is needed by Volume 2. Because the pool's soft capacity does not exceed its hard capacity, there is no way to allocate soft capacity to effectively "overcommit" the available hard capacity.

The final reserved space within the regular storage pool that is shown in Figure A-2 on page 427 is dedicated for snapshot use. The diagram shows that the specified snapshot reserve capacity of 34 GB is deducted from both the hard and soft space defined for the regular storage pool. This deduction guarantees that this space is available for consumption collectively by the snapshots associated with the pool. Although snapshots consume space at the partition level, the snapshot reserve capacity is still defined in increments of 17 GB.

The remaining 17 GB within the regular storage pool has not been allocated to either volumes or snapshots. All soft capacity remaining in the pool is “backed” by hard capacity. The remaining unused soft capacity is always less than or equal to the remaining unused hard capacity.

Thinly provisioned storage pool conceptual example

The thinly provisioned storage pool that was introduced in Figure A-1 on page 426 is explored in detail in Figure A-3. The hard capacity and the soft capacity allocated to this pool are the same in both diagrams: 136 GB of soft capacity and 85 GB of hard capacity are allocated. Because the available soft capacity exceeds the available hard capacity by 51 GB, you can thinly provision the volumes collectively by *up to* 66.7%, assuming that the snapshots are preserved and the remaining capacity within the pool is allocated to volumes.

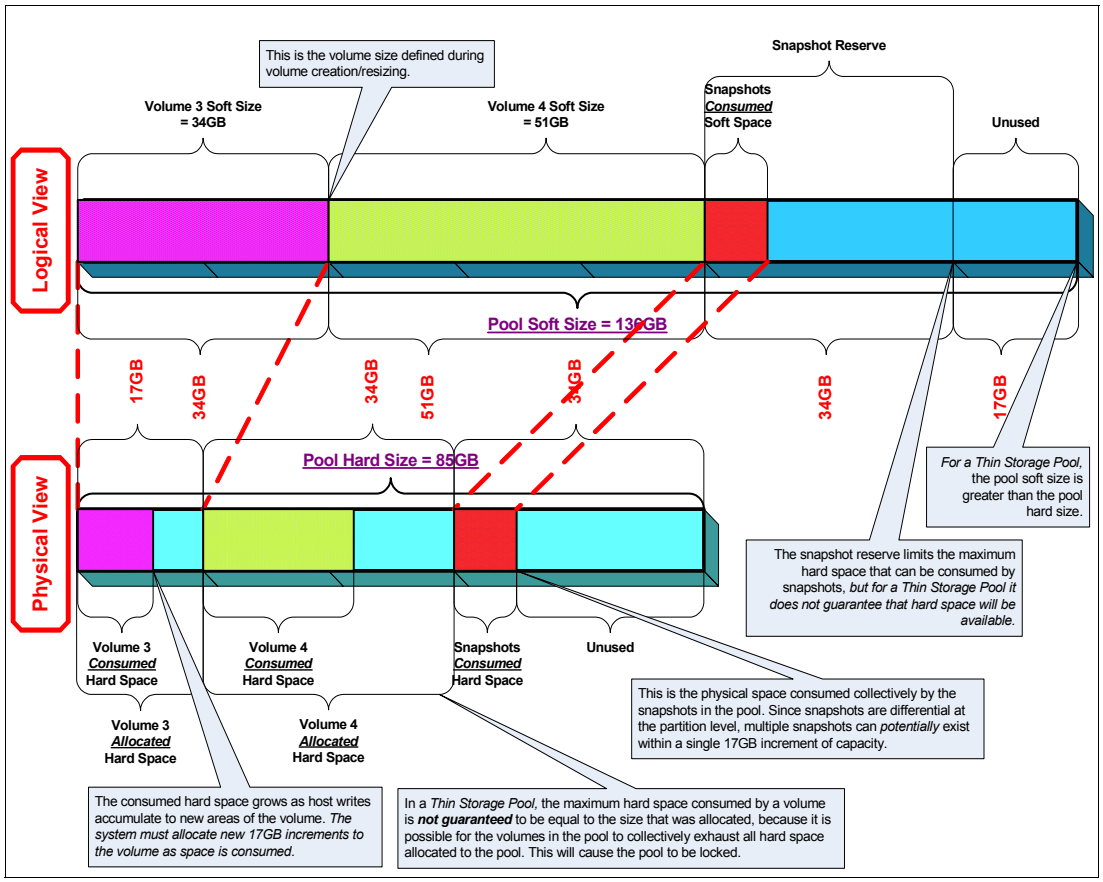

This is the volume size defined during volume creation/resizing.
Snapshot Reserve
Volume 3 Soft Size = 34GB
Volume 4 Soft Size = 51GB
Snapshots Consumed Soft Space
Unused
Logical View
Pool Soft Size = 136GB
17GB
34GB
34GB
51GB
34GB
34GB
17GB
Pool Hard Size = 85GB
Physical View
For a Thin Storage Pool, the pool soft size is greater than the pool hard size.
The snapshot reserve limits the maximum hard space that can be consumed by snapshots, but for a Thin Storage Pool it does not guarantee that hard space will be available.
Volume 3 Consumed Hard Space
Volume 4 Consumed Hard Space
Snapshots Consumed Hard Space
Unused
This is the physical space consumed collectively by the snapshots in the pool. Since snapshots are differential at the partition level, multiple snapshots can potentially exist within a single 17GB increment of capacity.
Volume 3 Allocated Hard Space
Volume 4 Allocated Hard Space
The consumed hard space grows as host writes accumulate to new areas of the volume. The system must allocate new 17GB increments to the volume as space is consumed.
In a Thin Storage Pool, the maximum hard space consumed by a volume is not guaranteed to be equal to the size that was allocated, because it is possible for the volumes in the pool to collectively exhaust all hard space allocated to the pool. This will cause the pool to be locked.

Figure A-3 Volumes and snapshot reserve space within a thinly provisioned storage pool

Consider Volume 3 in Figure A-3 on page 428. The size of the volume is defined as 34 GB. However, less than 17 GB has been consumed by host writes, so only 17 GB of hard capacity have been allocated by the system. In comparison, Volume 4 is defined as 51 GB, but Volume 4 has consumed 17 - 34 GB of hard capacity and therefore has been allocated 34 GB of hard space by the system. It is possible for either of these two volumes to require up to an additional 17 GB of hard capacity to become fully provisioned. Therefore, at least 34 GB of additional hard capacity must be allocated to this pool in anticipation of this requirement.

Finally, consider the 34 GB of snapshot reserve space shown in Figure A-3 on page 428. If a new volume is defined in the unused 17 GB of soft space in the pool, or if either Volume 3 or Volume 4 requires additional capacity, the system sacrifices the snapshot reserve space to give priority to the volume requirements. Normally, this scenario does not occur because additional hard space must be allocated to the storage pool as the hard capacity use crosses certain thresholds.

B

Additional LDAP information

In this appendix, we cover the following topics:

- Creating user accounts in Microsoft Active Directory
- Creating user accounts in Oracle Java Directory
- Securing Lightweight Directory Access Protocol (LDAP) communication with Secure Sockets Layer (SSL), including Windows Server SSL configuration and Oracle Java Directory SSL configuration
- Certificate authority setup

© Copyright IBM Corp. 2014. All rights reserved.

Creating user accounts in Microsoft Active Directory

Creating an account in Microsoft Active Directory for use by XIV Storage System LDAP authentication is similar to creating a regular user account. The only exception is the designated “description” attribute (field). This field must be populated with the predefined value for the authentication process to work.

Complete the following steps:

1. Start Active Directory Users and Computer by clicking **Start** → **Administrative Tools** → **Active Directory Users and Computers**.
2. Right-click the **Users** container and click **New** → **User**. The New Object - User window opens (Figure B-1).

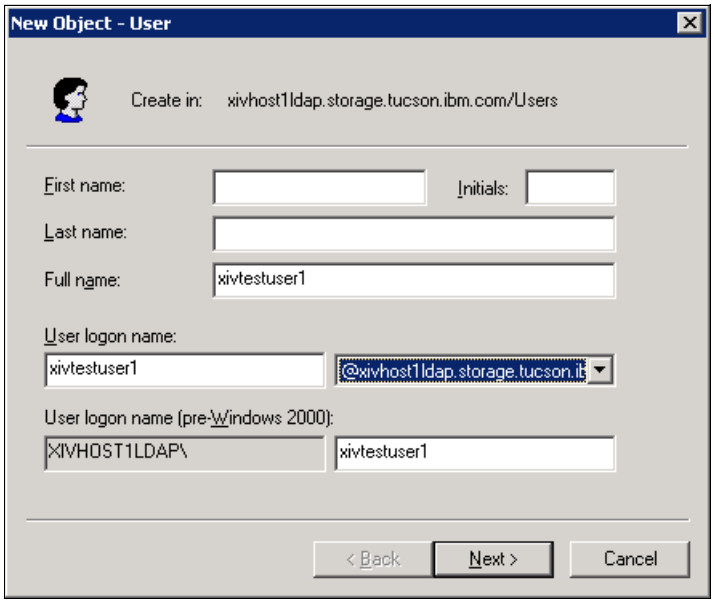


Figure B-1 Creating an Active Directory user account

The value entered in “Full name” is what the XIV Storage System uses as the user name. The only other mandatory field in this form is “User logon name”. The same `xivtestuser1` value is entered into both fields. The other fields can also be populated but are not required.

3. Create the account by clicking **Next**. A new window opens (Figure B-2).

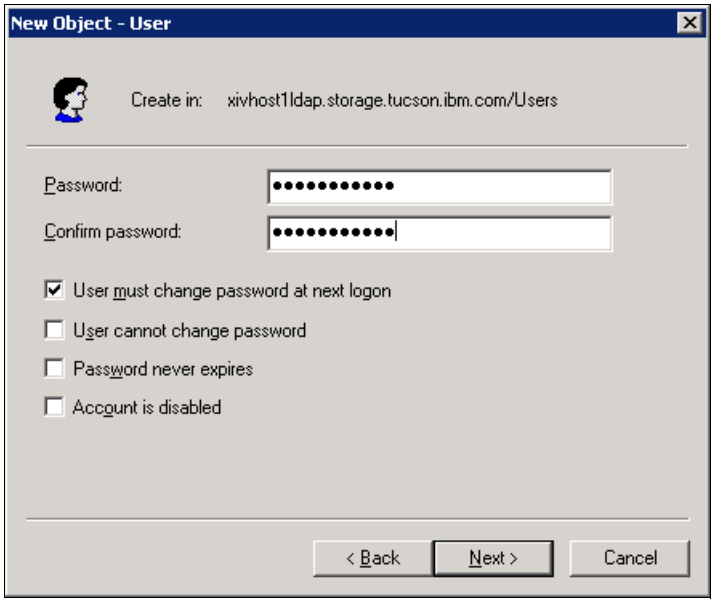


Figure B-2 Assigning a password

By default, the password is set to “User must change password at next login”. After the account is created, the user must log on to a server that is part of the Active Directory managed domain to change the password. After the password is changed, all the security rules and policies related to password management are in effect, such as password expiration, maintaining password change history, verifying password complexity, and so on.

Password: If the password initially assigned to an Active Directory user is not changed, XIV Storage System does not authenticate that user.

4. Complete the account creation by clicking **Next** → **Finish**.

5. Populate the Description field with a predefined value for the XIV Storage System category (role) mapping by selecting the `xivtestuser1` user name, right-clicking it, and clicking **Properties** (Figure B-3).

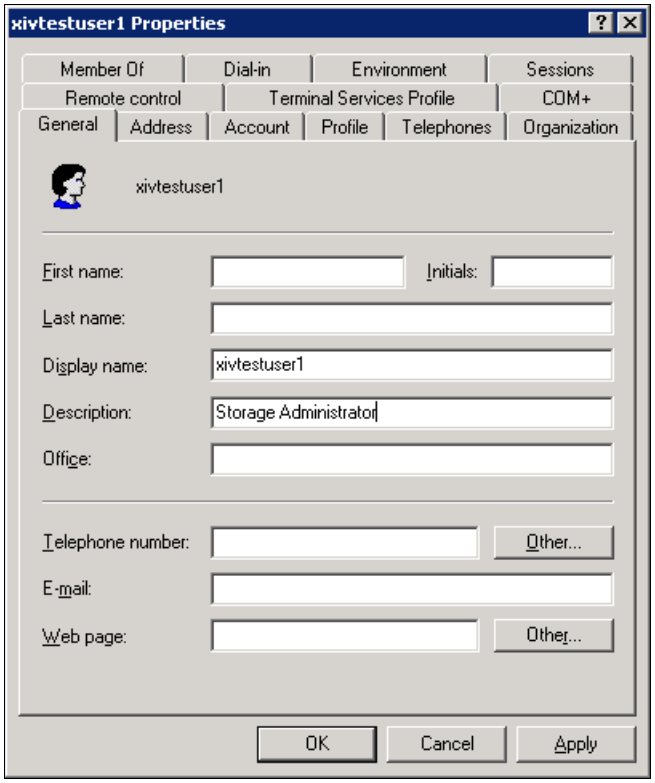


Figure B-3 Entering a predefined value into the Description field

Complete the account information update by clicking **OK**.

After the user account is created in Active Directory, its accessibility can be verified from any of the available LDAP clients. In our case, we used the OpenLDAP client, as shown in Example B-1.

Example: B-1 Active Directory account verification using OpenLDAP client

```
$ ldapsearch -x  -H "ldap://xivhost1.xivhost1ldap.storage.tucson.ibm.com:389"  -D
'CN=xivtestuser1,CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com' -w
pass2remember -b 'CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com'
cn=xivtestuser1

dn: CN=xivtestuser1,CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com
objectClass: top
objectClass: person
objectClass: organizationalPerson
objectClass: user
cn: xivtestuser1
description: Storage Administrator
distinguishedName: CN=xivtestuser1,CN=Users,DC=xivhost1ldap,DC=storage,DC=tucs
 on,DC=ibm,DC=com
instanceType: 4
whenCreated: 20090622172440.0Z
whenChanged: 20090622180134.0Z
```

```
displayName: xivtestuser1
uSNCreated: 98467
uSNChanged: 98496
name: xivtestuser1
objectGUID:: apHajqyazEyALYHDAJrjNA==
userAccountControl: 512
badPwdCount: 0
codePage: 0
countryCode: 0
badPasswordTime: 128901682350000000
lastLogoff: 0
lastLogon: 128901682415312500
pwdLastSet: 128901672940468750
primaryGroupID: 513
objectSid:: AQUAAAAAAAUVAAAAn59TxndIlskwvBQmdAQAAA==
accountExpires: 9223372036854775807
logonCount: 3
sAMAccountName: xivtestuser1
sAMAccountType: 805306368
userPrincipalName: xivtestuser1@xivhost1ldap.storage.tucson.ibm.com
objectCategory: CN=Person,CN=Schema,CN=Configuration,DC=xivhost1ldap,DC=storag
 e,DC=tucson,DC=ibm,DC=com
```

The **ldapsearch** command syntax might appear overly complex and its output seems difficult to interpret. However, this output might be the easiest way to verify that the account was created as expected. The **ldapsearch** command can also be useful for troubleshooting purposes when you are unable to communicate with an Active Directory LDAP server.

Here is a brief explanation of the **ldapsearch** command-line parameters:

- **-H 'ldap://xivhost1.xivhost1ldap.storage.tucson.ibm.com:389'**: Specifies that the LDAP search query must be sent to the "xivhost1.xivhost1ldap.storage.tucson.ibm.com" server using port number 389.
- **-D 'CN=xivtestuser1,CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com'**: The query is issued on behalf of the "xivtestuser1" user registered in "Users" container in the "xivhost1ldap.storage.tucson.ibm.com" Active Directory domain.
- **-w pass2remember**: The current password of the user "xivtestuser1" (after the initially assigned password was changed to this new password).
- **-b 'CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com'**: Base_DN, which is the location in the directory where to perform the search, that is, the "Users" container in the "xivhost1ldap.storage.tucson.ibm.com" Active Directory domain.
- **cn=xivtestuser1**: Specifies what object to search for.

The output of the **ldapsearch** command shows the structure of the LDAP object retrieved from the LDAP repository. We do not need to describe every attribute of the retrieved object, but at least two attributes must be checked to validate the response:

- name: xivtestuser1
- description: Storage Administrator

The fact that **ldapsearch** returns the expected results in our example indicates the following:

- The account is indeed registered in Active Directory.
- The distinguished name (DN) of the LDAP object is known and valid.

- The password is valid.
- The designated attribute "description" has a predefined value assigned "Storage Administrator".

When the Active Directory account verification is completed, we can proceed with configuring the XIV Storage System for LDAP authentication mode. We still have a few unassigned LDAP-related configuration parameters in our XIV Storage System, as shown in Example B-2.

Example: B-2 Remaining XIV LDAP configuration parameters

```
>> ldap_config_get
Name                      Value
base_dn
xiv_group_attrib          description
third_expiration_event    7
version                   3
user_id_attrib            objectSiD
current_server
use_ssl                   no
session_cache_period
second_expiration_event   14
read_only_role            Read Only
storage_admin_role        Storage Administrator
first_expiration_event    30
bind_time_limit           0
```

The parameters are defined as follows:

base_dn The base DN (distinguished name), which is the parameter that specifies where in the Active Directory LDAP repository that a user can be located. In our example, we use "CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com" as the base DN (see Example B-1 on page 434).

current_server A read-only parameter that cannot be populated manually. This parameter is updated by the XIV Storage System after the initial contact with the LDAP server is established.

session_cache_period Duration in minutes that the XIV Storage System keeps user credentials in its cache before discarding the cache contents. If a user repeats the login attempt within session_cache_period minutes from the first attempt, authentication is done from the cache content without contacting the LDAP server for the user credentials.

bind_time_limit The timeout value in seconds after which the next LDAP server on ldap_list_servers is called. The default value for this parameter is 0. It must be set to a non-zero value for bind (establishing an LDAP connection) to work. The rule also applies to configurations where the XIV Storage System is configured with only a single server on the ldap_list_servers list.

The populated values are shown in Example B-3.

Example: B-3 Completing and verifying LDAP configuration on XIV

```
>> ldap_config_set
base_dn="CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com"
session_cache_period=10 bind_time_limit=30
Command executed successfully.

$ xcli -c "XIV MN00019" -u ITSO -p redb00k  ldap_config_get
Name                       Value
base_dn CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com
xiv_group_attrib           description
third_expiration_event     7
version                    3
user_id_attrib             objectSiD
current_server
use_ssl                    no
session_cache_period       10
second_expiration_event    14
read_only_role             Read Only
storage_admin_role         Storage Administrator
first_expiration_event     30
bind_time_limit            30
```

To complete our description of the LDAP-related configuration parameters (at the XIV Storage System), we must describe the parameters that had default values assigned and did not need to be set explicitly:

version Version of the LDAP protocol used (the default is 3). This parameter must never be changed. Both Active Directory and Oracle Java Services Directory Server Enterprise Edition support LDAP protocol Version 3.

user_id_attrib LDAP attribute set to identify the user (in addition to user name) when recording user operations in the XIV event log. The default objectSiD value corresponds to the existing attribute name in Active Directory LDAP object class.

use_ssl Indicates if secure (SSL encrypted) LDAP communication is mandated. The default value is no. If set to yes without configuring both sides for SSL encrypted communication, LDAP authentication fails at the XIV Storage System.

first_expiration_event Number of days before the expiration of the certificate to set the first alert (severity “warning”). Set this parameter to several days that gives you enough time to generate and deploy a new security certificate.

second_expiration_event Number of days before the expiration of the certificate to set the second alert (severity “warning”).

third_expiration_event Number of days before the expiration of the certificate to set the third alert (severity “warning”).

Now that all configuration and verification steps are completed, the XIV Storage System is ready for the LDAP mode to be activated.

Creating user accounts in Oracle Java Directory

Creating an account in Oracle Java Directory can be done in many ways using various LDAP clients. We used the LDAP graphical user interface (GUI) client that is part of the Java System Directory Service Control Center web tool that is part of the Oracle Java Directory Server product suite.

The designated `description` attribute must be populated with the predefined value for the authentication process to work. From the Oracle Java Directory LDAP Server perspective, assigning a value to the `description` attribute is not mandatory and is not enforced by the server itself. An LDAP server allows the creation of an account with no value assigned to the attribute. However, this attribute value is required by the XIV Storage System to establish LDAP role mapping. This field must be populated with the predefined value for the authentication process to work.

To start the Oracle Java System Directory Service Control Center, point your browser to the IP address of your Oracle Java Directory LDAP Server for a secure connection on port 6789. In our example, we use the following URL to access the Oracle Java System Directory Service Control Center:

“`https://xivhost2.storage.tucson.ibm.com:6789`**”**

Before the first user account can be created, the LDAP administrator must create a suffix. A suffix (also known as a *naming context*) is a DN that identifies the top entry in the directory hierarchy. A Oracle Java Directory LDAP server can have multiple suffixes, each identifying a locally held directory hierarchy, for example, `o=ibm`, or, in our specific example, `dc=xivauth`.

To create a suffix, log on to the Java Console using your own user ID and password and select the **Directory Service Control Center (DSCC)** link in the Services section. Authenticate to the Directory Service Manager application. In the Common Tasks tab, click **Directory Entry Management** → **Create New Suffix or Replication Topology**. Then, complete the following steps:

1. Enter Suffix Name: Specify the new suffix DN. In our example, we use `dc=xivauth`. Click **Next**.
2. Choose Replication Options: Accept the default **Do Not Replicate Suffix** (LDAP replication is beyond the scope for this book).
3. Choose Servers: Click **xivhost2.storage.tucson.ibm.com:389** in the Available Servers list and click **Add**. The server name is displayed in the Chosen Servers list. Click **Next**.
4. Choose Settings: Accept the default of **Use Default Settings**.
5. Choose Database Location Options. Accept the default database location. Click **Next**.
6. Choose Data Options: Select **Create Top Entry for the Suffix**. Click **Next**.
7. Review the settings for the suffix about to be created and click **Finish** if they are correct.

After the new suffix creation is confirmed, you can proceed with LDAP entry creation.

To create an LDAP entry login to the Java Console using your own user ID and password, select the **Directory Service Control Center (DSCC)** link in the Services section. Authenticate to Directory Service Manager application. In the Common Tasks tab, click **Directory Entry Management** → **Create New Entry**.

Now you can start the New Entry configuration wizard by completing the following steps:

1. Select a server instance. Oracle Java Directory allows you to create multiple instances of an LDAP server. However, the only instance that uses port 389 for non-SSL LDAP and port 636 for SSL LDAP communication can be used for XIV LDAP authentication services. Select an instance configured on port 384, as shown in Figure B-4.

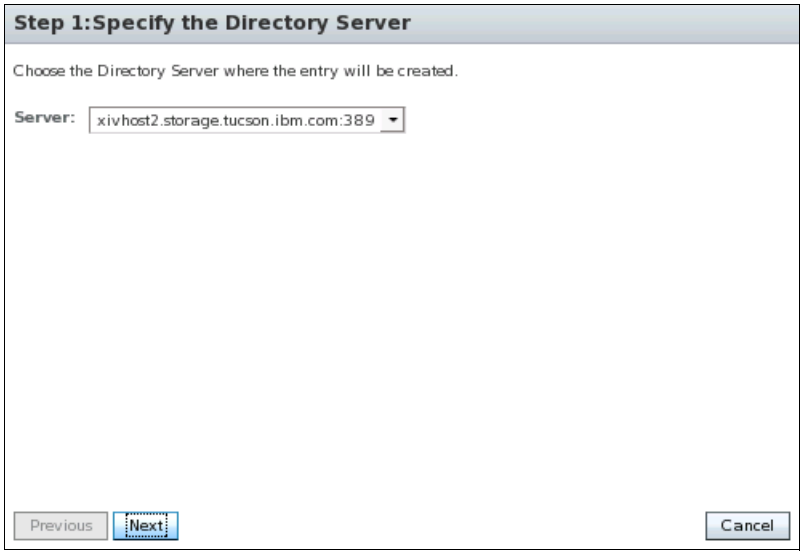


Figure B-4 Selecting a Directory Server instance

2. Select the new entry location, as shown in Figure B-5. The LDAP administrator determines the location of a new entry. Unlike the Active Directory LDAP repository, where location is directly linked to the domain name, Oracle Java Directory LDAP server provides greater flexibility in terms of placement for the new entry.

Figure B-5 Selecting the entry location

The location of all entries for XIV Storage System accounts must be the same because the XIV Storage System has only one LDAP configuration parameter that specifies the location. In our example, we use `dc=xivauth` as the entry location for XIV Storage System user accounts. The location name is the same as the suffix name in this example. There are certain similarities between Windows file system and LDAP directory structures. You can think of LDAP suffixes as drive letters. A drive letter can contain directories, but you can also put files onto the root directory on a drive letter. In our example, we put a new account at the level (by analogy) of a root directory, the `dc=xivauth` location.

As your LDAP repository grows, it might no longer be practical to put types of entries into the same location. In this case, just like with Windows file system, you can create subdirectories and place new entries there. The LDAP equivalent of what has become a directory hierarchy in your file system is called the *Directory Information Tree* (*DIT*). After the entry location is selected and the XIV Storage System is configured to point to that location, all the new account entries can be created only in that location.

3. Select an object class for the new entry. Unlike a predefined object class for a user account in Active Directory LDAP, Oracle Java Directory LDAP presents you with a choice of object class types.

   The LDAP object class describes the content and purpose of the object. It also contains a list of attributes, such as a name, surname, or telephone number. Traditionally, the `inetOrgPerson` object class type is used for LDAP objects that describe personal information. To be compatible with XIV Storage System, an object class must include a minimal set of attributes. These attributes are as follows:

`uid`	User identifier (user name)
`userPassword`	User password
`description` (configurable)	LDAP role mapping attribute

   You can select another object class type if it contains the same minimal set of attributes. The object class type can enforce certain rules. For example, various attributes can be designated as mandatory, in which case a new LDAP object cannot be created without assigning a value to that attribute. For the `inetOrgPerson` object, there are two mandatory attributes:

`cn`	Full Name, also called Common Name
`sn`	Full Name, also called Surname

Although it is possible to populate these two objects with various values, we use the `uid` value to populate both the `cn` and `sn` attributes. See Figure B-6.

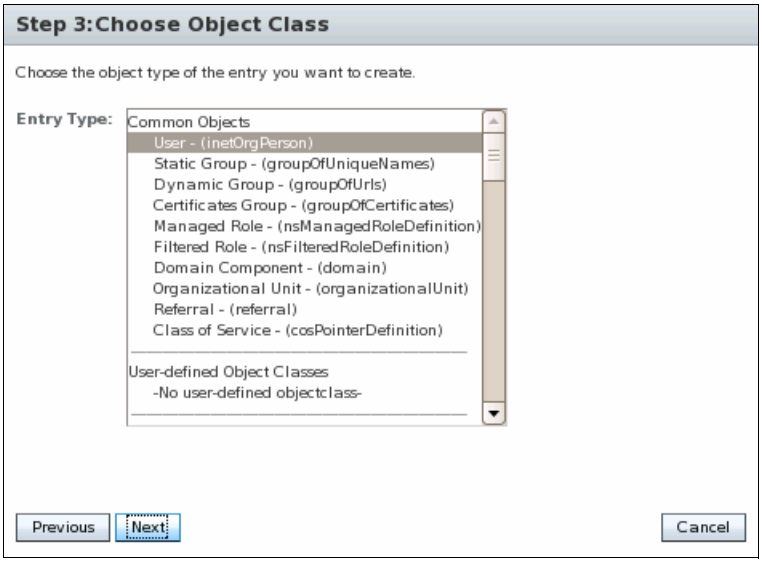


Figure B-6 Object class selection

4. Enter the attribute values. The first field, Naming Attribute, must remain **uid**; XIV Storage System uses that attribute name for account identification. We populate the mandatory attributes with values. You can also choose to populate other optional attributes and store their values in the LDAP repository, but XIV Storage System does not use those attributes. See Figure B-7 on page 442.

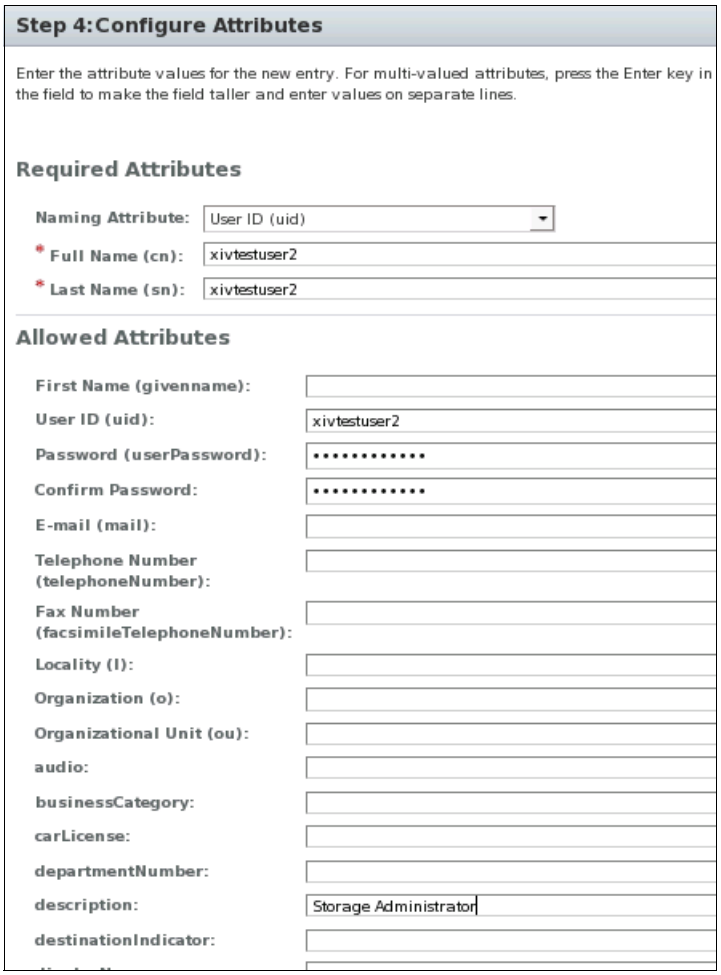


Figure B-7 Entering object attribute values

5. A Summary window (Figure B-8) shows what you have selected and entered in the previous steps. You can go back and change parameters if you want to do so. Otherwise, proceed with the entry creation.

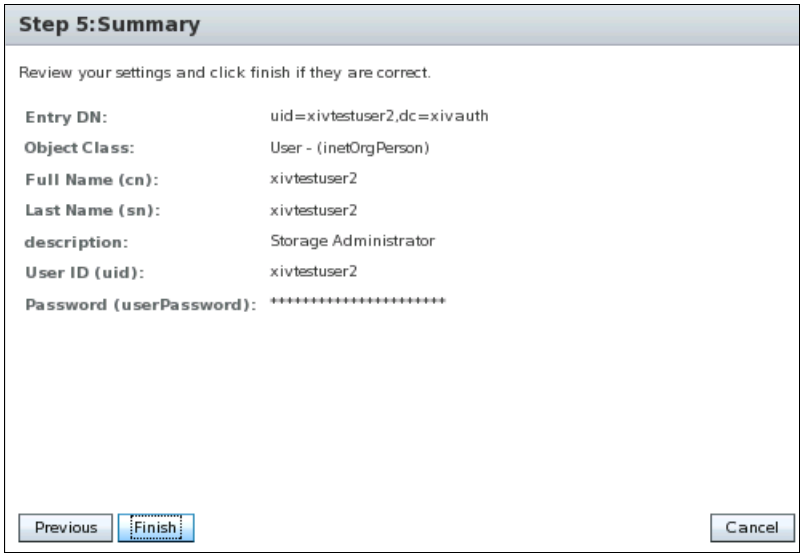


Figure B-8 Reviewing entry settings

If all the information was entered correctly, you get an Operation Completed Successfully message in the window that opens (Figure B-9). If the operation failed for one reason or another, you must go back and make necessary changes before resubmitting your request.

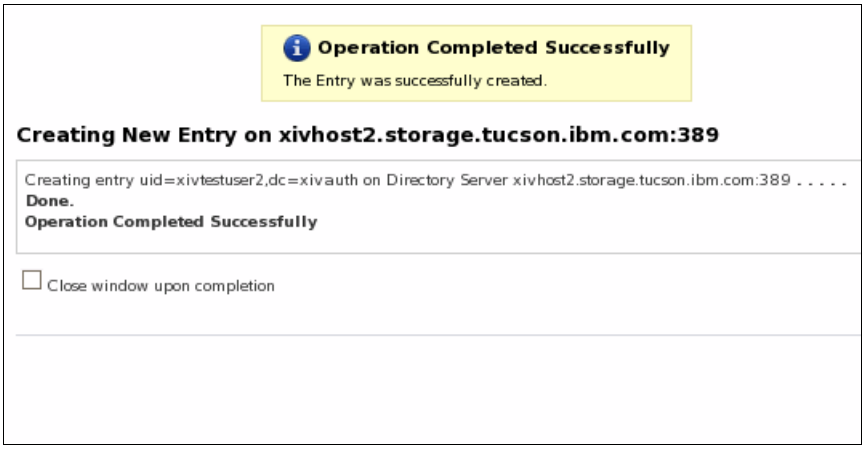

Operation Completed Successfully
The Entry was successfully created.
Creating New Entry on xivhost2.storage.tucson.ibm.com:389
Creating entry uid=xivtestuser2,dc=xivauth on Directory Server xivhost2.storage.tucson.ibm.com:389 . . . . .
Done.
Operation Completed Successfully
Close window upon completion

Figure B-9 Entry creation confirmation

After the user account is created in Oracle Java Directory LDAP, its accessibility can be verified by using any of the available LDAP clients. In our example (Example B-4 on page 444), we use the Oracle Java Directory LDAP client.

Example: B-4 Oracle Java Directory account verification using Oracle Java Directory LDAP client

```
$ /opt/sun/dsee6/bin/ldapsearch -b dc=xivauth -h xivhost2.storage.tucson.ibm.com
-D uid=xivtestuser2,dc=xivauth -w pwd2remember uid=xivtestuser2

dn: uid=xivtestuser2,dc=xivauth
uid: xivtestuser2
description: Storage Administrator
objectClass: inetOrgPerson
objectClass: organizationalPerson
objectClass: person
objectClass: top
sn: xivtestuser2
cn: xivtestuser2
```

The **ldapsearch** command syntax might appear overly complex and its output seems difficult to interpret. However, this output might be the easiest way to verify that the account was created as expected. The **ldapsearch** command can also be useful for troubleshooting purposes when you are unable to communicate with a Oracle Java Directory LDAP server.

Here is a brief explanation of the **ldapsearch** command-line parameters:

-h xivhost2.storage.tucson.ibm.com	Specifies that the LDAP search query must be sent to the `xivhost2.xivhost1ldap.storage.tucson.ibm.com` server using the default port 389.
-b dc=xivauth	`Base_DN`, which is the location in the DIT.
-D uid=xivtestuser2,dc=xivauth	The query issued on behalf of the `xivtestuser2` user in the `dc=xivauth` Oracle Java Directory repository.
-w pwd2remember	The current password of the `xivtestuser2` user.
uid=xivtestuser2	Specifies what object to search.

The output of the **ldapsearch** command shows the structure of the object found. We do not need to describe every attribute of the returned object, but at least two attributes must be checked to validate the response:

- `uid: xivtestuser2`
- `description: Storage Administrator`

The fact that **ldapsearch** returns the expected results in our example indicates the following:

1. The account is registered in Oracle Java Directory.
2. We know where in the Oracle Java Directory repository the account is located.
3. We know the valid password.
4. The designated attribute `description` has a predefined value of `Storage Administrator`.

When the Oracle Java Directory account verification is complete, we proceed with configuring XIV Storage System for LDAP authentication mode. We still have a few unassigned LDAP-related configuration parameters in our XIV Storage System, as shown in Example B-5 on page 445.

Example: B-5 Remaining XIV LDAP configuration parameters

```
>> ldap_config_get
Name                      Value
base_dn
xiv_group_attrib          description
third_expiration_event    7
version                   3
user_id_attrib            objectSiD
current_server
use_ssl                   no
session_cache_period
second_expiration_event   14
read_only_role            Read Only
storage_admin_role        Storage Administrator
first_expiration_event    30
bind_time_limit           0
```

The parameters are as follows:

base_dn Base DN, which is the parameter that specifies where in the Oracle Java Directory DIT a user can be located. In our example, we use `dc=xivauth` as base DN.

user_id_attrib `The` LDAP attribute set to identify the user (in addition to the user name) when recording user operations in the XIV Storage System event log. The default value for the attribute is `objectSiD`, which is suitable for Active Directory but not for Oracle Java Directory LDAP. The `objectSiD` attribute is not defined in the `inetOrgPerson` object class used by Oracle Java Directory. In our example, we set it to `uid`.

current_server A read-only parameter that cannot be populated manually. It is updated by the XIV Storage System after the initial contact with the LDAP server is established.

session_cache_period The duration in minutes that the XIV Storage System keeps user credentials in its cache before discarding the cache contents. If a user repeats a login attempt within session_cache_period minutes from the first attempt, authentication is done based on the cache content without contacting the LDAP server for user credentials.

bind_time_limit The timeout value in seconds after which the next LDAP server on the `ldap_list_servers` is called. The default value for this parameter is `0`. It must be set to a non-zero value for bind (establishing LDAP connection) to work. The rule also applies to configurations where the XIV Storage System is configured with only a single server on the `ldap_list_servers` list.

The populated values are shown in Example B-6.

Example: B-6 Completing and verifying LDAP configuration on XIV

```
$ xcli -c "ARCXIVJEMT1" -u admin -p s8cur8pwd ldap_config_set base_dn="dc=xivauth"
user_id_attrib=uid session_cache_period=10 bind_time_limit=30
Command executed successfully.

$ xcli -c "XIV MN00019" -u admin -p s8cur8pwd ldap_config_get
Name                      Value
base_dn                    dc=xivauth
xiv_group_attrib           description
third_expiration_event     7
version                    3
user_id_attrib             sid
current_server
use_ssl                    no
session_cache_period       10
second_expiration_event    14
read_only_role             Read Only
storage_admin_role         Storage Administrator
first_expiration_event     30
bind_time_limit            30
```

To complete our description of the LDAP-related configuration parameters (at the XIV Storage System), we describe the parameters that had default values assigned and did not need to be set explicitly:

version Version of LDAP protocol used (the default is 3). This parameter must never be changed. Both Active Directory and Oracle Java Services Directory Server Enterprise Edition support LDAP protocol Version 3.

user_id_attrib LDAP attribute set to identify the user (in addition to user name) when recording user operations in the XIV Storage System event log.

use_ssl Indicates if secure (SSL encrypted) LDAP communication is mandated. If set to yes without configuring both sides for SSL, any encrypted communication for LDAP authentication on the XIV Storage System fails.

first_expiration_event Number of days before the expiration of the certificate to set the first alert (severity warning). Set this parameter to several days that gives you enough time to generate and deploy a new security certificate.

second_expiration_event Number of days before the expiration of the certificate to set the second alert (severity warning).

third_expiration_event Number of days before the expiration of the certificate to set the third alert (severity warning).

Securing LDAP communication with SSL

In any authentication scenario, information is exchanged between the LDAP server and XIV Storage System where access is being sought. Secure Sockets Layer (SSL) can be used to implement secure communications between the LDAP client and server. LDAP over SSL (LDAPS), the secure version of LDAP protocol, allows secure communication between the XIV Storage System and LDAP server by using encrypted SSL connections. This protocol allows a setup where user passwords never appear in clear text.

SSL provides methods for establishing identity using X.509 certificates and ensuring message privacy and integrity using encryption. To create an SSL connection, the LDAP server must have a digital certificate signed by a trusted certificate authority (CA). Companies have the choice of using a trusted CA from a vendor or creating their own certificate authority. In this scenario, the `xivauth.org` CA is used.

To be operational, SSL must be configured on both the client and the server. Server configuration includes generating a certificate request, obtaining a server certificate from a CA, and installing the server and CA certificates.

Windows Server SSL configuration

To configure SSL for LDAP on a Windows Server, you must install the Management Console (MMC) snap-in to manage local certificates. Then, you must create a certificate request (CER) and have the CER signed by a CA. Finally, you import the signed certificate into the local keystore, import a CA certificate as a trusted root CA, and then reboot the server for the new configuration to take effect.

Installation of the local certificates Management Console snap-in

Install the certificate snap-in for MMC to allow you to manage the certificates in your local machine keystore by completing the following steps:

1. Start the MMC by selecting **Start** → **Run**. Enter mmc /a and select **OK**.
2. Click the **File** → **Add/Remove Snap-in** menu to open the Add/Remove Snap-in window.
3. Click **Add** to open the Add Standalone Snap-In window. Click the **Certificates** snap-in and then click **Add**.
4. Select the **Computer Account** option to manage system-wide certificates. Click **Next** to continue.
5. Click the **Local Computer** option to manage certificates on the local computer only. Click **Finish** and then **Close**. Then, click **OK** to complete the snap-in installation.
6. Click **File** → **Save as** and save the console configuration in the `%SYSTEMROOT%\system32` directory with a file name of `localcert.msc`.
7. Create a shortcut in the Administrative Tools folder in your Start menu by right-clicking the **Start** menu and then **Open All Users**. Click the **Program** folder and then the **Administrative Tools** folder.
8. Click **File** → **New** → **Shortcut**. Enter the location of the saved console, `%SYSTEMROOT%\system32\localcert.msc`, in the “Type the location of the item” field. Then, click **Next** to continue.
9. Enter the name of the new shortcut, Certificates (Local Computer), in the “Type a name for this shortcut” field.
10. To start the local certificate management tool, click **Start** → **Administrative tools** → **Certificates** (**Local Computer)**.

When the local certificate management tool starts, it appears as shown in Figure B-10. The certificates used by Active Directory are in the **Console Root** → **Certificates (Local Computer)** → **Personal** → **Certificates** folder. The list of trusted root certificates authorities is in the **Console Root** → **Certificates (Local Computer)** → **Trusted Certification Authorities** → **Certificates** folder.

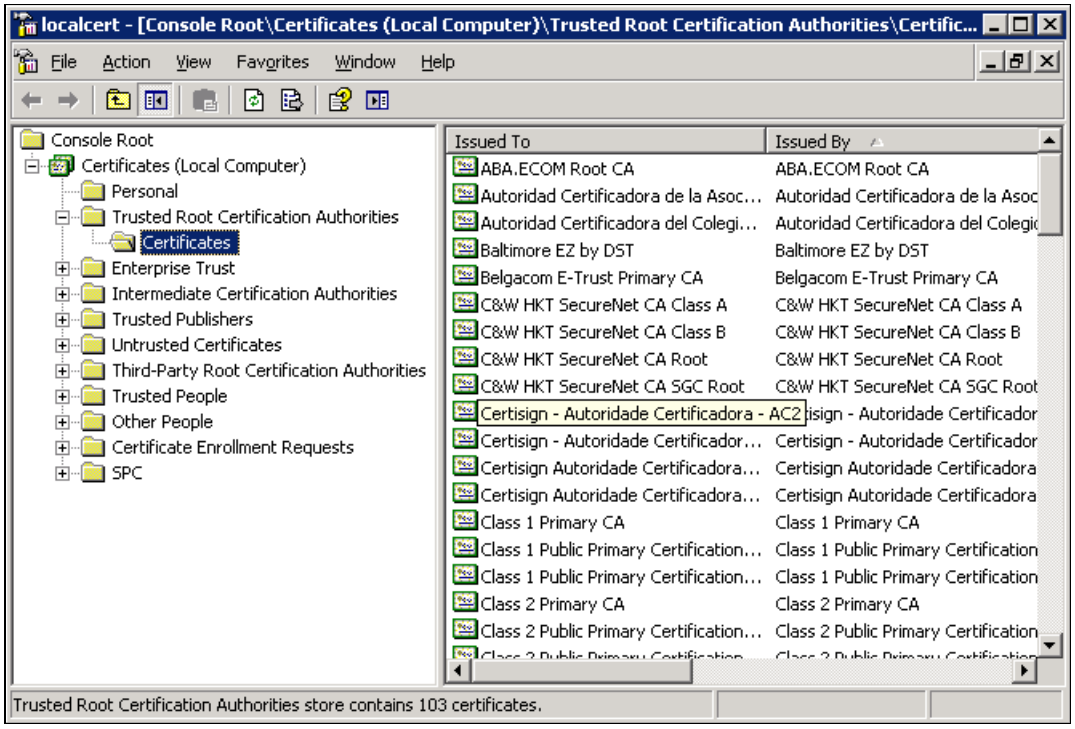


Figure B-10 Windows certificate MMC snap-in (local computer)

Generating a Windows Server certificate request

You must use the **`certreq`** command to generate a certificate request. The **`certreq`** command uses a text instruction file, which specifies the attributes needed to generate a certificate. It contains attributes such as the subject's common name, certificate key length, and additional key usage extensions. Active Directory requires that the certificate meet the following requirements:

- The private key and certificate for the local machine must be imported into the local computer's personal keystore.
- The fully qualified domain name (FQDN) for the Active Directory must be in the common name (CN) in the subject field or DNS entry in the subject alternative name extension.
- The certificate must be issued by a CA that the Active Directory server and the XIV Storage System trust.
- The certificate must contain the enhanced key usage extension that specifies the server authentication object identifier (OID) `1.3.6.1.5.5.7.3.1`. This OID indicates that the certificate is used as an SSL server certificate.

Example B-7 shows the text instruction file used to generate the certificate for the xivhost1ldap.storage.tucson.ibm.com domain controller. The subject field is set to CN=xivhost1.xivhost1ldap.storage.tucson.ibm.com, which is the FQDN of the domain controller. Use the **certreq** command to generate the certificate request file.

Example: B-7 Text instruction file for certificate request generation

```
[Version]

Signature="$Windows NT$

[NewRequest]

Subject = "CN=xivhost1.xivhost1ldap.storage.tucson.ibm.com"
KeySpec = 1
KeyLength = 1024
; Can be 1024, 2048, 4096, 8192, or 16384.
; Larger key sizes are more secure, but have
; a greater impact on performance.
Exportable = TRUE
MachineKeySet = TRUE
SMIME = False
PrivateKeyArchive = FALSE
UserProtected = FALSE
UseExistingKeySet = FALSE
ProviderName = "Microsoft RSA SChannel Cryptographic Provider"
ProviderType = 12
RequestType = PKCS10
KeyUsage = 0xa0

[EnhancedKeyUsageExtension]

OID=1.3.6.1.5.5.7.3.1 ; this is for Server Authentication

C:\SSL\> certreq -new xivhost1_cert_req.inf xivhost1_cert_req.pem
C:\SSL\>
```

Signing and importing a Windows server certificate

After the CER is generated (`xivhost1_cert_req.pem`), you must send the request to the certificate authority to be signed. For more information about signing this certificate, see “Signing a certificate for the xivhost1 server” on page 462. After the signed certificate is returned, you must import the certificate into the local machine’s personal keystore.

Example B-8 shows how to import the signed certificate by running the **`certreq`** command. Confirm that the certificate is imported correctly by running **`certutil`**.

Example: B-8 Accepting the signed certificate into the local certificate keystore

```
C:\>certreq -accept xivhost1_cert.pem
C:\SSL>certutil -store my
================ Certificate 0 ================
Serial Number: 01
Issuer: E=ca@xivstorage.org, CN=xivstorage, O=xivstorage, L=Tucson, S=Arizona,
C=US
Subject: CN=xivhost1.xivhost1ldap.storage.tucson.ibm.com
Non-root Certificate
Cert Hash(sha1): e2 8a dd cc 84 47 bc 49 85 e2 31 cc e3 23 32 c0 ec d2 65 3a
  Key Container =
227151f702e7d7b2105f4d2ce0f6f38e_8aa08b0a-e9a6-4a73-9dce-c84e45aec165
  Provider = Microsoft RSA SChannel Cryptographic Provider
Encryption test passed
CertUtil: -store command completed successfully.
```

Importing a certificate authority certificate

Until the `xivstorage.org` CA is designated as a trusted root, any certificate signed by that CA is untrusted. You must import the CA’s certificate, using the local certificate management tool, into the Trusted Certification Authorities folder in the local keystore.

To start the local certificate management tool, click **Start** → **Administrative tools** → **Certificates** (**Local Computer)** and complete the following steps:

1. After the certificate tool opens, select the `/Console Root/Certificates (Local Computer)/Trusted Certification Authorities` folder.
2. Start the certificate import wizard by clicking **Action** → **All Tasks** → **Import**. Click **Next** to continue.
3. Select the file that you want to import. The `xivstorage.org` CA certificate is in the `cacert.pem` file. Click **Next** to continue.
4. Select the **Place all certificates in the following store** option and ensure that the certificate store field is set to **Trusted Root Certification Authorities**. Click **Next** to continue.
5. The CA certificate is now imported. Click **Finish** to close the wizard.

After the CA and server certificates are imported into the local keystore, you can then use the local certificate management tool to check whether the certificates are correctly imported. Open the **Console Root** → **Certificates (Local Computer)** → **Personal** → **Certificates** folder and select the certificate issued to `xivhost1.xivhost1ldap.storage.tucson.ibm.com`.

Figure B-11 shows that the certificate that was issued to `xivhost1.xivhost1ldap.storage.tucson.ibm.com` is valid and was issued by the xivstorage CA. The certificate has a corresponding private key in the keystore. The “Ensures the identity of the remote computer” text indicates that the certificate has the required server authentication key usage defined.

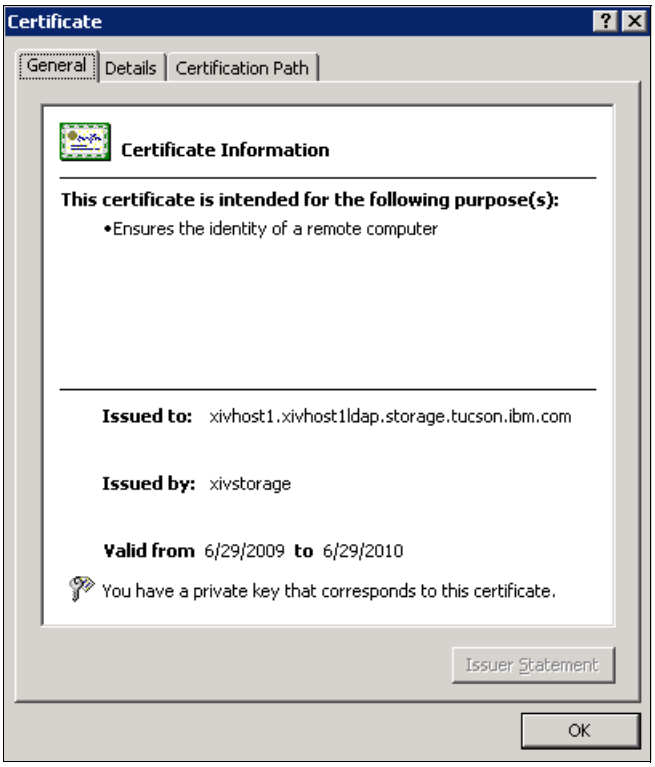


Figure B-11 Certificate information window

To check the xivstorage certificate, open the **Console Root** → **Certificates (Local Computer)** → **Trusted Certification Authorities** → **Certificates** folder and select the certificate issued by xivstorage. Figure B-12 shows that the certificate issued to and by the xivstorage CA is valid.

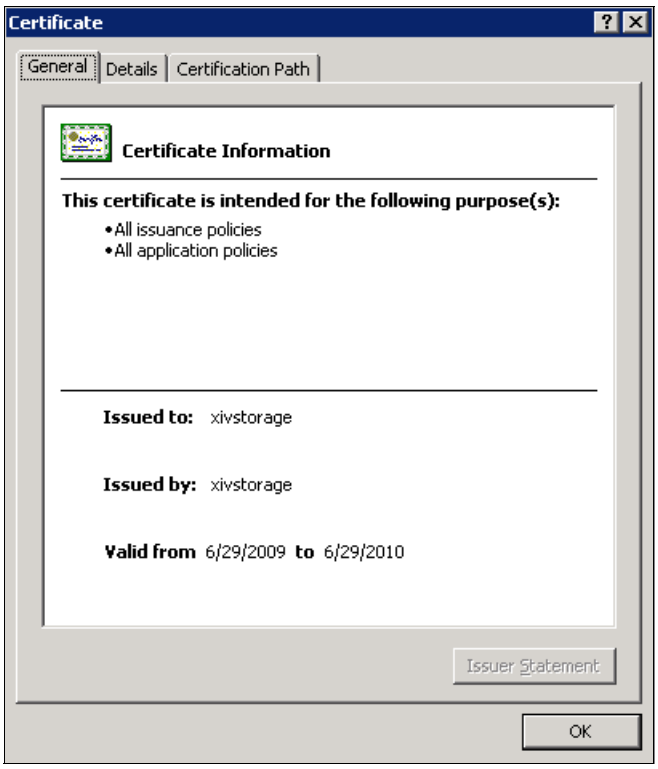


Figure B-12 Certificate information window for xivstorage certificate authority

Low-level SSL validation using the openssl command

The easiest way to test the low-level SSL connection to the LDAP server is by running the **`openssl s_client`** command with the `-showcerts` option. This command connects to the specified host and lists the server certificate, the certificate authority chain, supported ciphers, SSL session information, and verify return code. If the SSL connection worked, the **`openssl s_client`** command result in the verify return code is `0 (ok)`.

Example B-9 shows the output of the **`openssl s_client`** command connecting a Linux server (`xivstorage.org`) to the Active Directory server (`xivhost1.xivhost1ldap.storage.tucson.ibm.com`). This command connects to the Active Directory server using the secure LDAP port (636).

Example: B-9 Low-level SSL validation using the openssl s_client

```
openssl s_client  -host xivhost1.xivhost1ldap.storage.tucson.ibm.com -port 636
-CAfile cacert.pem  -showcerts
   ...
   ---
Server certificate
subject=/CN=xivhost1.xivhost1ldap.storage.tucson.ibm.com
issuer=/C=US/ST=Arizona/L=Tucson/O=xivstorage/CN=xivstorage/emailAddress=ca@xivsto
rage.org
---
   ...
New, TLSv1/SSLv3, Cipher is RC4-MD5
```

```
Server public key is 1024 bit
SSL-Session:
    Protocol  : TLSv1
    Cipher    : RC4-MD5
    Session-ID: 9E240000CE9499A4641F421F523ACC347ADB91B3F6D3ADD5F91E271B933B3F4F
    Session-ID-ctx:
    Master-Key:
F05884E22B42FC4957682772E8FB1CA7772B8E4212104C28FA234F10135D88AE496187447313149F2E
8922OE6F4DADF3
    Key-Arg   : None
    Krb5 Principal: None
    Start Time: 1246314540
    Timeout   : 300 (sec)
    Verify return code: 0 (ok)
```

Attention: To complete the configuration of SSL for the Active Directory, you must reboot the Windows server.

Basic secure LDAP validation using the ldapsearch command

After you have confirmed that the SSL connection is working, you must confirm that you are able to search your LDAP directory using LDAP on a secure port. This action confirms that the LDAP server can communicate using an SSL connection.

In our example, we use the OpenLDAP client for SSL connection validation. A CA certificate needs to be added to the key ring file used by the OpenLDAP client. The TLS_CERTS option in the OpenLDAP configuration file (typically, `/etc/openldap/ldap.conf`) specifies the file that contains certificates for all of the certificate authorities the client recognizes. See Example B-10.

Example: B-10 Testing LDAP over SSL by using the ldapsearch command

```
# /usr/bin/ldapsearch -x  -H
“ldaps://xivhost1.xivhost1ldap.storage.tucson.ibm.com:636” -D
'CN=xivtestuser1,CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com' -w
pass2remember -b 'CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com'

dn: CN=xivtestuser1,CN=Users,DC=xivhost1ldap,DC=storage,DC=tucson,DC=ibm,DC=com
objectClass: top
objectClass: person
objectClass: organizationalPerson
objectClass: user
cn: xivtestuser1
description: Storage Administrator
distinguishedName: CN=xivtestuser1,CN=Users,DC=xivhost1ldap,DC=storage,DC=tucs
 on,DC=ibm,DC=com
...
# search result
search: 2
result: 0 Success
```

The Uniform Resource Identifier (URI) format used with the **-H** option specifies that LDAPS must be used on port 636 (the LDAP secure port).

Oracle Java Directory SSL configuration

This section describes the use of an SSL protocol for communicating with the Oracle Java Directory.

Creating a Oracle Java Directory certificate request

To configure SSL for Oracle Java LDAP Directory, you must create a CER, and have the CER signed by a CA. Then, import the signed certificate into the local keystore, import a CA certificate as a trusted root CA, and then restart the LDAP server for the new configuration to take effect.

Generating a Oracle Java Directory server certificate request

To generate a certificate request using the Oracle Java web console tool, complete the following steps:

1. Point your web browser to HTTPS port 6789, in our example, "`https://xivhost2.storage.tucson.ibm.com:6789`".
2. Log on to the system and select **Directory Service Control Center (DSCC)** and authenticate to Directory Service Manager.
3. Click **Directory Servers** → **xivhost2.storage.tucson.ibm.com:389** → **Security** → **Certificates** → **Request CA-Signed Certificate**. Complete the certificate request form. A sample of the certificate request form is shown in Figure B-13.

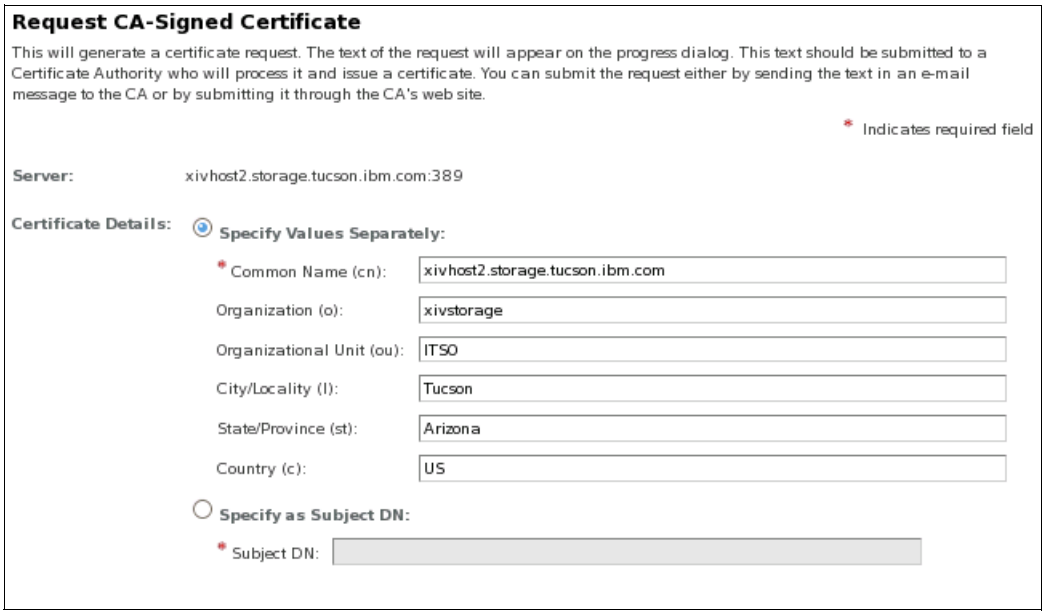

Request CA-Signed Certificate
This will generate a certificate request. The text of the request will appear on the progress dialog. This text should be submitted to a Certificate Authority who will process it and issue a certificate. You can submit the request either by sending the text in an e-mail message to the CA or by submitting it through the CA's web site.
* Indicates required field
Server:
xivhost2.storage.tucson.ibm.com:389
Certificate Details:
Specify Values Separately:
* Common Name (cn):
xivhost2.storage.tucson.ibm.com
Organization (o):
xivstorage
Organizational Unit (ou):
ITSO
City/Locality (l):
Tucson
State/Province (st):
Arizona
Country (c):
US
Specify as Subject DN:
* Subject DN:

Figure B-13 Certificate request

4. Copy the generated certificate shown in Figure B-14. Request into the `xivhost2_cert_req.pem` file.

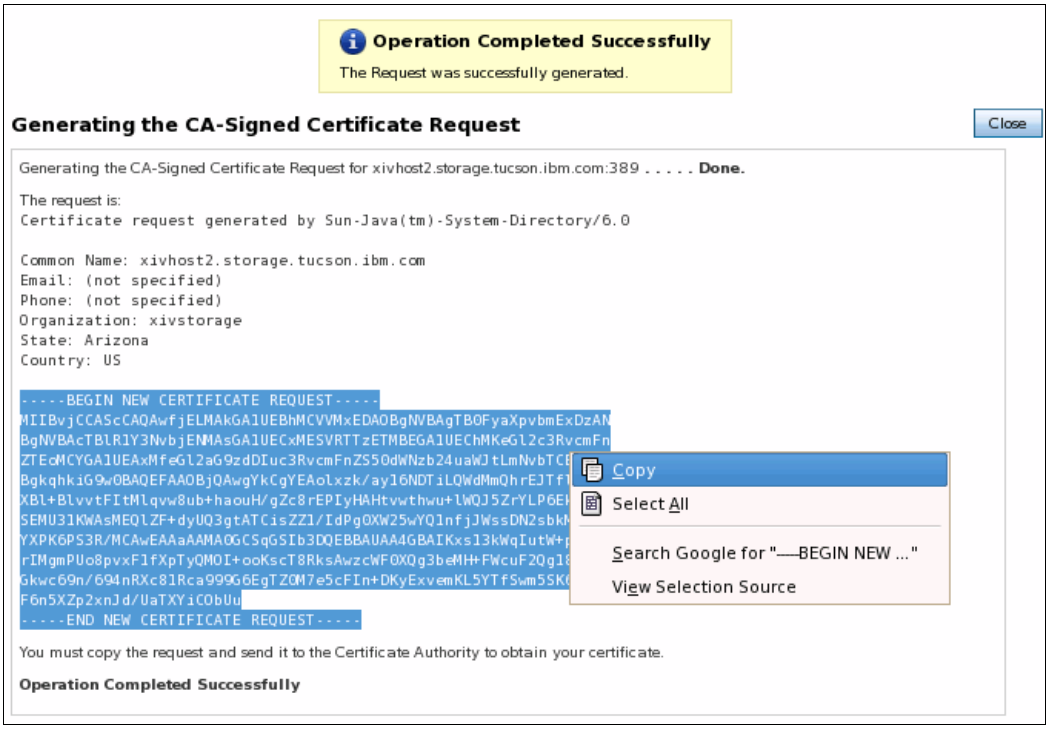


Figure B-14 Generated certificate

Signing and importing a server certificate

After the CER is generated (`xivhost2_cert_req.pem`), you must send the request to the certificate authority to be signed. For more information about signing this certificate, see “Signing a certificate for the xivhost1 server” on page 462. After the signed certificate `xivhost2_cert.pem` file is returned, you must import the certificate into the local machine’s personal keystore.

To add the signed certificate to the “Directory Service Manager” manager application, click **Directory Servers** → **xivhost2.storage.tucson.ibm.com:389** → **Security** → **Certificates** → **Add**. Copy and paste the certificate stored in the `xivhost2_cert.pem` file, as shown in Figure B-15.

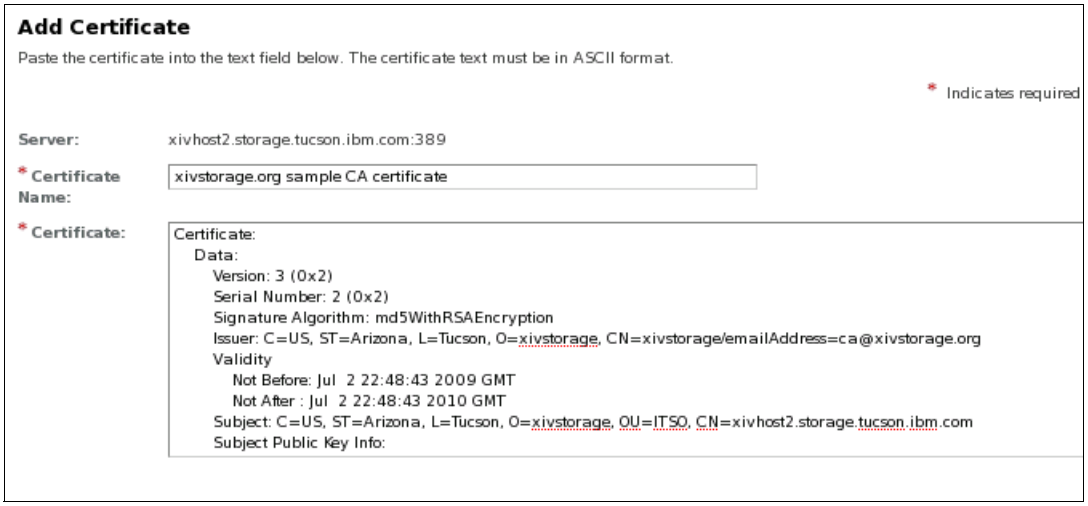

Add Certificate
Paste the certificate into the text field below. The certificate text must be in ASCII format.
* Indicates required
Server:
xivhost2.storage.tucson.ibm.com:389
* Certificate Name:
xivstorage.org sample CA certificate
* Certificate:
Certificate:
Data:
Version: 3 (0x2)
Serial Number: 2 (0x2)
Signature Algorithm: md5WithRSAEncryption
Issuer: C=US, ST=Arizona, L=Tucson, O=xivstorage, CN=xivstorage/emailAddress=ca@xivstorage.org
Validity
Not Before: Jul 2 22:48:43 2009 GMT
Not After : Jul 2 22:48:43 2010 GMT
Subject: C=US, ST=Arizona, L=Tucson, O=xivstorage, OU=ITSO, CN=xivhost2.storage.tucson.ibm.com
Subject Public Key Info:

Figure B-15 Adding signed certificate

Importing a certificate authority certificate

Until the `xivstorage.org` CA is designated as a trusted root, any certificate signed by that CA is untrusted. You must import the CA’s certificate using the “Directory Service Manager” manager application by completing the following steps:

1. Click **Directory Servers** → **xivhost2.storage.tucson.ibm.com:389** → **Security** → **CA Certificates** → **Add**.
2. Copy and paste the certificate authority certificate stored in the `cacert.pem` file, as shown in Figure B-16.

Figure B-16 Importing the certificate authority certificate

3. After the CA and signed certificates are imported into the local keystore, you can use the local certificate management tool to check whether the certificates are correctly imported. Click **Directory Servers** → **xivhost2.storage.tucson.ibm.com:389** → **Security** → **Certificates** and click the **xivstorage.org sample CA certificate** link.

Figure B-17 shows that the certificate issued to `xivhost2.storage.tucson.ibm.com` is valid and was issued by the xivstorage certificate authority.

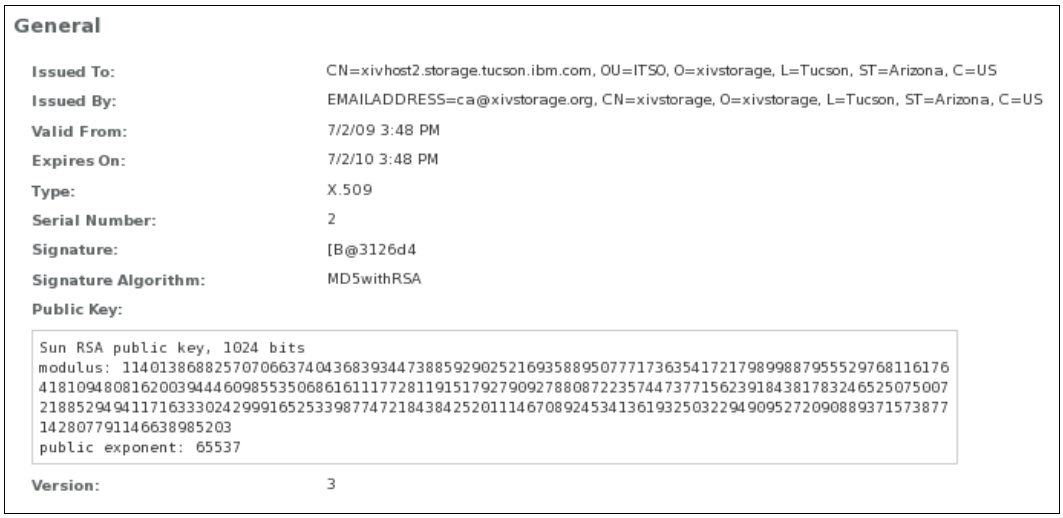


Figure B-17 Signed Oracle Java Directory certificate information

4. To check the xivstorage certificate, click **Directory Servers** → **xivhost2.storage.tucson.ibm.com:389** → **Security** → **CA Certificates** and click the **xivstorage.org sample CA Certificate Authority certificate** link. Figure B-18 shows that the certificate issued to and by the xivstorage CA is valid.

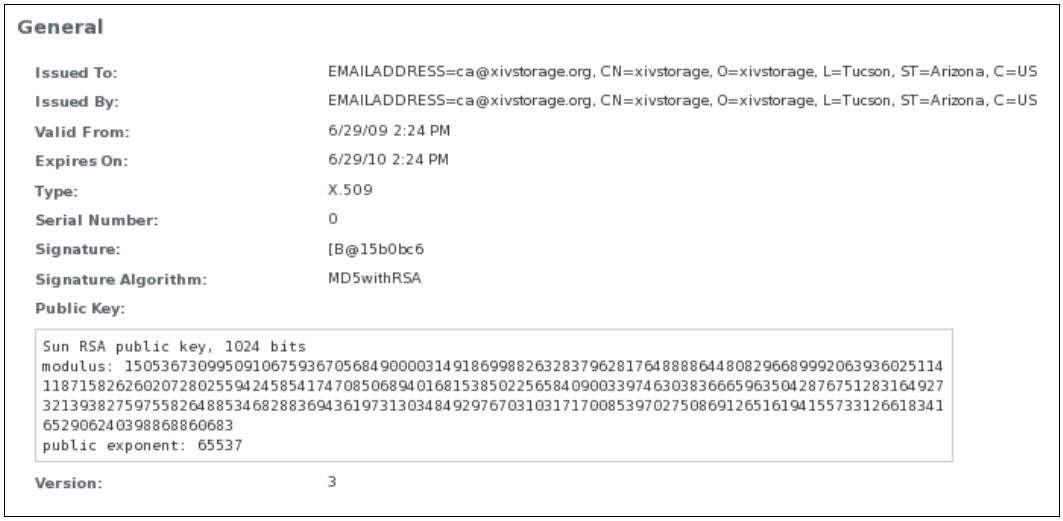


Figure B-18 Certificate information for xivstorage certificate authority

5. To activate the imported certificate, click **Directory Servers** → **xivhost2.storage.tucson.ibm.com:389** → **Security**. In the General tab, click the **Certificate** drop-down menu and select **xivstorage.org sample CA certificate**, as shown in Figure B-19 on page 458.

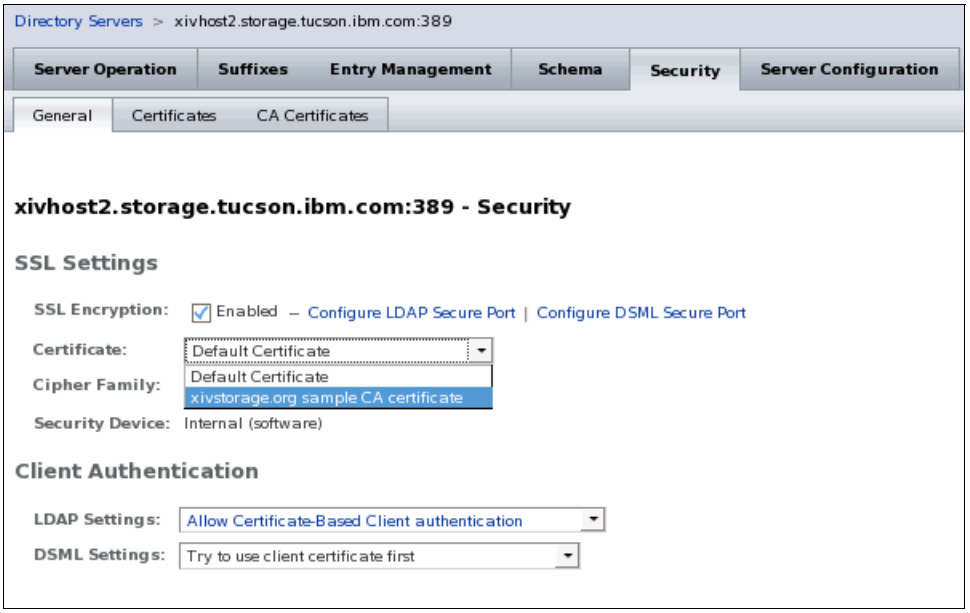


Figure B-19 New signed certificate activation

As shown in Figure B-20, you are prompted to restart the LDAP server for the new certificate to take effect.

Restart Required
You have modified the Certificate used by the Directory Server. You must restart the Directory Server in order this modification to be taken into account.
Note: Once the Java(TM) Web Console session has ended this message will disappear.

Figure B-20 Manual restart request after activating new certificate

Low-level SSL validation using the openssl command

The easiest way to test the low-level SSL connection to the LDAP server is by running the **`openssl s_client`** command with the **`-showcerts`** option. This command connects to the specified host and lists the server certificate, the certificate authority chain, supported ciphers, SSL session information, and verify return code. If the SSL connection works, the **`openssl s_client`** command result in the verify return code is `0 (ok)`.

Example B-11 shows the output of the **`openssl s_client`** command connecting a Linux server (`xivstorage.org`) to the Oracle Java Directory server `xivhost2.storage.tucson.ibm.com`. This command connects to the Oracle Java Directory server using the secure LDAP port (636).

Example: B-11 Low-level SSL validation using the openssl s_client

```
# openssl s_client  -host xivhost2.storage.tucson.ibm.com -port 636 -CAfile
cacert.pem  -showcerts
   ...
---
Server certificate
subject=/C=US/ST=Arizona/L=Tucson/O=xivstorage/OU=ITSO/CN=xivhost2.storage.tucson.
ibm.com
issuer=/C=US/ST=Arizona/L=Tucson/O=xivstorage/CN=xivstorage/emailAddress=ca@xivsto
rage.org
---
```

```
Acceptable client certificate CA names
/O=Sun Microsystems/CN=Directory Server/CN=636/CN=xivhost2.storage.tucson.ibm.com
/C=US/ST=Arizona/L=Tucson/O=xivstorage/CN=xivstorage/emailAddress=ca@xivstorage.or
g
---
SSL handshake has read 2144 bytes and written 328 bytes
---
New, TLSv1/SSLv3, Cipher is AES256-SHA
Server public key is 1024 bit
Compression: NONE
Expansion: NONE
SSL-Session:
    Protocol  : TLSv1
    Cipher    : AES256-SHA
    Session-ID: 48B43B5C985FE1F6BE3F455F8350A4155DD3330E6BD09070DDCB80DCCB570A2E
    Session-ID-ctx:
    Master-Key:
1074DC7ECDD9FC302781C876B3101C9C618BB074020D7062E7EA3AB794CA9C5D1A33447EE254288CEC
86BBB6CD264DCA
    Key-Arg   : None
    Krb5 Principal: None
    Start Time: 1246579854
    Timeout   : 300 (sec)
    Verify return code: 0 (ok)
---
```

Basic secure LDAP validation using the ldapsearch command

After you have confirmed that the SSL connection is working properly, you must verify that you are able to search your LDAP directory using LDAPS on port 636. This action confirms that the LDAP server can communicate using an SSL connection.

In Example B-12, we use OpenLDAP client for the SSL connection validation. A CA certificate needs to be added to the key ring file used by OpenLDAP client. The TLS_CERTS option in the OpenLDAP configuration file (typically, `/etc/openldap/ldap.conf`) specifies the file that contains certificates for all of the certificate authorities that the client recognizes.

Example: B-12 Testing LDAP over SSL by using the ldapsearch command

```
# /usr/bin/ldapsearch -x  -H “ldaps://xivhost2.storage.tucson.ibm.com:636” -D
'uid=xivtestuser2,dc=xivauth' -w pwd2remember -b 'dc=xivauth'

# extended LDIF
#
# LDAPv3
# base <dc=xivauth> with scope subtree
# filter: uid=xivtestuser2
# requesting: ALL
#

# xivtestuser2, xivauth
dn: uid=xivtestuser2,dc=xivauth
uid: xivtestuser2
objectClass: inetOrgPerson
objectClass: organizationalPerson
objectClass: person
```

```
objectClass: top
sn: xivtestuser2
cn: xivtestuser2
description: custom_role_01

# search result
search: 2
result: 0 Success
```

The URI format used with the **-H** option specifies that LDAPS must be used on port 636 (the LDAP secure port).

Certificate authority setup

This section describes the setup and use of the certificate authority that was used with all example scenarios in this book to issue certificates.

OpenSSL comes with most Linux distributions by default. Information about OpenSSL can be found at the OpenSSL website:

http://www.openssl.org

Creating the CA certificate

To set up the CA for the `xivstorage.org` domain, we need to make certain assumptions. We modify `openssl.cnf` to reflect these assumptions to the CA. That file can be found in `/usr/share/ssl/openssl.cnf`. Portions of that file are shown in Example B-13.

Example: B-13 openssl.cnf

```
[ CA_default ]

dir             = /root/xivstorage.orgCA# Where everything is kept
certs           = $dir/certs            # Where the issued certs are kept
crl_dir         = $dir/crl              # Where the issued crl are kept
database        = $dir/index.txt        # database index file.
new_certs_dir   = $dir/newcerts         # default place for new certs.
certificate     = $dir/cacert.pem       # The CA certificate
serial          = $dir/serial           # The current serial number
crl             = $dir/crl.pem          # The current CRL
private_key     = $dir/private/cakey.pem# The private key
RANDFILE        = $dir/private/.rand    # private random number file
x509_extensions = usr_cert              # The extentions to add to the cert
name_opt        = ca_default            # Subject Name options
cert_opt        = ca_default            # Certificate field options
default_days    = 365                   # how long to certify for
default_crl_days= 30                    # how long before next CRL
default_md      = md5                   # which md to use.
preserve        = no                    # keep passed DN ordering
copy_extensions = copy                   # Extension copying option
default_days    = 365                   # how long to certify for
default_crl_days= 30                    # how long before next CRL
default_md      = md5                   # which md to use.
preserve        = no                    # keep passed DN ordering
   .
```

```
   .
   .
[ req_distinguished_name ]
countryName                     = Country Name (2 letter code)
countryName_default             = US
countryName_min                 = 2
countryName_max                 = 2

stateOrProvinceName             = State or Province Name (full name)
stateOrProvinceName_default     = TX
localityName                    = Locality Name (eg, city)
localityName_default            = Tucson
O.organizationName              = Organization Name (eg, company)
O.organizationName_default      = xivstorage
organizationalUnitName          = Organizational Unit Name (eg, section)
commonName                      = xivstorage.org (eg, your server's hostname)
commonName_max                  = 64
emailAddress                    = ca@xivstorage.org
emailAddress_max                = 64
```

The directories to store the certificates and keys must be created by running the following command:

```
mkdir /root/xivstorage.orgCA /root/xivstorage.orgCA/certs
/root/xivstorage.orgCA/crl /root/xivstorage.orgCA/newcerts
/root/xivstorage.orgCA/private
```

OpenSSL uses a couple of files to maintain the CA. These files must be created by running the following commands:

- **touch /root/xivstorage.orgCA/index.txt**
- **echo "01" >> /root/xivstorage.orgCA/serial**

The access rights on the directories and files need to be reviewed to restrict access to the CA and, most importantly, to the private key as far as possible.

To certify the CA certificate for 365 days, run the **OpenSSL** command directly, as shown in Example B-14.

Example: B-14 Generating the CA certificate

```
openssl req -new -x509 -days 365 -keyout /root/xivstorage.orgCA/private/cakey.pem
-out /root/xivstorage.orgCA/cacert.pem
Generating a 1024 bit RSA private key
....++++++
.....................++++++
writing new private key to '/root/xivstorage.orgCA/private/cakey.pem'
Enter PEM pass phrase:
Verifying - Enter PEM pass phrase:
-----
You are about to be asked to enter information that will be incorporated
into your certificate request.
What you are about to enter is what is called a Distinguished Name or a DN.
There are quite a few fields but you can leave some blank
For some fields there will be a default value,
If you enter '.', the field will be left blank.
-----
```

```
Country Name (2 letter code) [US]:
State or Province Name (full name) [Arizona]:
Locality Name (eg, city) [Tucson]:
Organization Name (eg, company) [xivstorage]:
Organizational Unit Name (eg, section) []:
Common Name (eg, YOUR name) []:xivstorage
Email Address []:ca@xivstorage.org
```

During the creation of the certificate, any missing information must be provided. Also, the information that has been defined by using the defaults in the `openssl.cnf` file must be confirmed. The password for the CA private key must be entered during the creation process. This password is needed whenever the CA's private key is used. The following command can be used to view the CA certificate:

```
openssl x509 -in cacert.pem -text
```

Signing a certificate

The client or server that needs to obtain a certificate must create a certificate signing request and send this request to the CA.

Certificate request details can be viewed by running the following command:

```
openssl req -in xivhost1_cert_req.pem -text
```

`xivhost1_cert_req.pem` is the certificate signing request, and is generated on the `xivhost1.xivhost1ldap.storage.tucson.ibm.com` server.

Signing a certificate for the xivhost1 server

To sign the certificate, run the **`openssl`** command with a specified policy, as shown in Example B-15.

Example: B-15 Signing certificate for xivhost1 server

```
# openssl ca -policy policy_anything -cert cacert.pem -keyfile private/cakey.pem
-out  xivhost1_cert.pem -in xivhost1_cert_req.pem
Using configuration from /usr/share/ssl/openssl.cnf
Enter pass phrase for private/cakey.pem:
Check that the request matches the signature
Signature ok
Certificate Details:
        Serial Number: 1 (0x1)
        Validity
            Not Before: Jun 29 21:35:33 2009 GMT
            Not After : Jun 29 21:35:33 2010 GMT
        Subject:
            commonName                =
xivhost1.xivhost1ldap.storage.tucson.ibm.com
        X509v3 extensions:
            X509v3 Basic Constraints:
            CA:FALSE
            Netscape Comment:
            OpenSSL Generated Certificate
            X509v3 Subject Key Identifier:
            C8:EB:8D:84:AB:86:BB:AF:5B:74:4D:35:34:0E:C5:84:30:A1:61:84
            X509v3 Authority Key Identifier:
            keyid:A8:0B:D1:B5:D6:BE:9E:61:62:E3:60:FF:3E:F2:BC:4D:79:FC:E3:5A
```

```
DirName:/C=US/ST=Arizona/L=Tucson/O=xivstorage/CN=xivstorage/emailAddress=ca@xivst
orage.org
            serial:00

            X509v3 Extended Key Usage:
            TLS Web Server Authentication
            X509v3 Key Usage:
            Digital Signature, Key Encipherment
Certificate is to be certified until Jun 29 21:35:33 2010 GMT (365 days)
Sign the certificate? [y/n]:y

1 out of 1 certificate requests certified, commit? [y/n]y
Write out database with 1 new entries
Data Base Updated
```

Related publications

The publications listed in this section are considered particularly suitable for more detailed information about the topics covered in this book.

IBM Redbooks publications

For information about ordering this publication, see “How to get IBM Redbooks publications” on page 466. The following publications might be available in softcopy only:

- *IBM XIV Storage System Business Continuity Functions*, SG24-7759
- *IBM XIV Storage System Multi-site Mirroring*, REDP-5129
- *XIV Storage System: Host Attachment and Interoperability*, SG24-7904
- *XIV Storage System SSD Caching Implementation*, REDP-4842
- *Using the IBM XIV Storage System with OpenStack Cloud Environments*, REDP-4971
- *XIV Storage System in VMware Environments*, REDP-4965
- *IBM Hyper-Scale for the XIV Storage System*, REDP-5053
- *RESTful API Support in XIV*, REDP-5064
- *XIV Security with Data-at-Rest Encryption*, REDP-5047
- *XIV Gen3 with SVC and Storwize V7000*, REDP-5063

Other publications

These publications are also relevant for further information:

- *IBM XIV Storage System Planning Guide,* SC27-5412
- *IBM XIV Storage System: Product Overview*, GC27-3912
- *IBM XIV Storage System User Manual*, GC27-3914
- *IBM XIV Remote Support Proxy Installation and User’s Guide*, GA32-0795
- *IBM XIV Storage System Application Programming Interface,* GC27-3916
- *IBM XIV Storage System Management Tools Operations Guide*, SC27-5986
- *IBM XIV Storage System XCLI Utility User Manual*, GC27-3915
- *IBM Hyper-Scale Manager 1.5 Installation as application*, GC27-5984
- *IBM Hyper-Scale Manager 1.5 Installation as a virtual appliance*, GC27-5985

Online resources

These sources are also useful for more information:

- IBM XIV Storage System documentation, IBM Knowledge Center

  http://ibm.co/1rvfciG
- IBM XIV Storage System web page

  http://www.ibm.com/systems/storage/disk/xiv/index.html
- IBM System Storage Interoperation Center (SSIC)

  http://www.ibm.com/systems/support/storage/ssic/interoperability.wss
- Storage Networking Industry Association (SNIA) website

  http://www.snia.org/

© Copyright IBM Corp. 2014. All rights reserved.

- IBM Systems Director Downloads web page

  http://www.ibm.com/systems/director/downloads/
- IBM Systems Director documentation, IBM Knowledge Center

  http://www.ibm.com/support/knowledgecenter/SSAV7B/welcome

How to get IBM Redbooks publications

You can search for, view, or download IBM Redbooks publications, Redpaper publications, Technotes, draft publications, and Additional materials, as well as order hardcopy IBM Redbooks publications, at this website:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Redbooks

IBM XIV Storage System Architecture and Implementation

(1.0” spine)
0.875”<->1.498”
460 <-> 788 pages

IBM XIV Storage System
Architecture and Implementation

Multi-tenancy gives more flexibility for XIV in Cloud environments

Enhanced performance classes allow virtual tiers of storage

6 TB drives offer lower cost per TB and lower power consumption

This IBM Redbooks publication describes the concepts, architecture, and implementation of the IBM XIV Storage System. The XIV is a scalable enterprise storage system that is based on a grid array of hardware components. It can attach to both Fibre Channel Protocol (FCP) and IP network Small Computer System Interface (iSCSI) capable hosts. This system is a good fit for clients who want to be able to grow capacity without managing multiple tiers of storage. The XIV is suited for mixed or random access workloads, including online transaction processing, video streaming, images, email, and emerging workload areas, such as Web 2.0 and cloud storage.

The focus of this edition is on the XIV Gen3 running Version 11.5 of the XIV system software, which brings enhanced value for the XIV Storage System in cloud environments. It offers multitenancy support, VMware vCloud Suite integration, more discrete performance classes, and RESTful API enhancements that expand cloud automation integration. Version 11.5 introduces support for three-site mirroring to provide high availability and disaster recovery. It also enables capacity planning through the Hyper-Scale Manager, mobile push notifications for real-time alerts, and enhanced security.

We describe many of the unique and powerful concepts that form the basis of the XIV Storage System logical and physical architecture and the planning and preparation tasks for deploying the system by using the XIV Storage Manager GUI or the XIV command-line interface. We also describe the performance characteristics options for alerting and monitoring, including enhanced secure remote support.

This book is for IT professionals who want an understanding of the XIV Storage System. It is also for readers who need detailed advice on how to configure and use the system.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-7659-08 ISBN 0738439096